

AD-A208 219

5

AFGL-TR-89-0024

**ANALYSIS OF THE NONLINEAR PARAMETRIC LEAST-SQUARES
ADJUSTMENT VIA AN ISOMORPHIC GEOMETRICAL SETUP
WITH TENSOR STRUCTURE**

Georges Blaha

Nova University Oceanographic Center
8000 North Ocean Drive
Dania, Florida 33004

Scientific Report No. 1

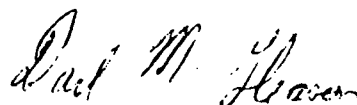
June 1988

Approved for public release; distribution unlimited

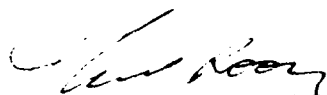
AIR FORCE GEOPHYSICS LABORATORY
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
HANSCOM AFB, MASSACHUSETTS 01731-5000

DTIC
ELECTRIC
MAY 18 1989
S E D

This technical report has been reviewed and is approved for publication.



DAVID M. GLEASON
Contract Manager



THOMAS P. ROONEY, Chief
Geodesy & Gravity Branch

FOR THE COMMANDER



DONALD H. ECKHARDT, Director
Earth Sciences Division

This report has been reviewed by the ESD Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS).

Qualified requestors may obtain additional copies from the Defense Technical Information Center. All others should apply to the National Technical Information Service.

If your address has changed, or if you wish to be removed from the mailing list, or if the addressee is no longer employed by your organization, please notify AFGL/DAA, Hanscom AFB, MA 01731-5000. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document requires that it be returned.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
AFGL-TR-89-0024		
4. TITLE (and Subtitle)		5. TYPE OF REPORT & PERIOD COVERED
ANALYSIS OF THE NONLINEAR PARAMETRIC LEAST-SQUARES ADJUSTMENT VIA AN ISOMORPHIC GEOMETRICAL SETUP WITH TENSOR STRUCTURE		Scientific Report No. 1
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s)		8. CONTRACT OR GRANT NUMBER(s)
Georges Blaha		F19628-87-K-0030
9. PERFORMING ORGANIZATION NAME AND ADDRESS		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
Nova University Oceanographic Center 8000 North Ocean Drive Dania, Florida 33004		61102F 2309G1RV
11. CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE
Air Force Geophysics Laboratory Hanscom AFB, Massachusetts 01731-5000 Contract Monitor: Mr. David M. Gleason/LWG		June 1988
		13. NUMBER OF PAGES
		56
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report)
		Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report)		
Approved for public release; distribution unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)		
Parametric Adjustment	Tensor analysis	Observational space
Nonlinear model	Metric tensor	Model surface
Least-squares criterion	Associated metric tensor	Model plane
Variance-covariances	Orthonormal vectors	Minimum distance
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)		
<p>The parametric adjustment model expresses n observables in terms of u parameters, where the structure linking the two groups is in general nonlinear. This model can be expanded in the Taylor series based on an initial set of parameters. The standard treatment includes only the first-order partial derivatives of the observables with respect to the parameters, constituting a linearized model which is resolved via normal equations as mandated by the least-squares criterion. If necessary, the solution is iterated upon using the same linearized model in conjunction with an updated set of parameters.</p>		

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 63 IS OBSOLETE

Unclassified

In this study, the resolution of a nonlinear adjustment model is addressed through an isomorphic geometrical setup with tensor structure and notation, represented by a u -dimensional "model surface" embedded in a flat n -dimensional "observational space". The n observations correspond to the observational-space coordinates of the point Q , the u initial parameters correspond to the model-surface coordinates of the "initial" point P , and the u adjusted parameters correspond to the model-surface coordinates of the "least-squares" point \bar{P} . The least-squares criterion results in a minimum-distance property implying that the vector $\bar{P}Q$ must be orthogonal to the model surface.

The geometrical setup leads to the solution of "modified normal equations" characterized by a positive-definite matrix. The latter contains the second-order and, optionally, the third-order partial derivatives of the observables with respect to the parameters. This approach significantly shortens the convergence process as compared to the standard (linearized) approach. The variance-covariance and the weight matrices of adjusted quantities correspond respectively to the associated metric tensors and to the metric tensors at \bar{P} , formulated in the pertinent Riemannian manifolds.

Unclassified

TABLE OF CONTENTS

	Page
1. INTRODUCTION	1
2. GEOMETRICAL SETUP	4
3. SOLUTION	8
3.1 General Development	8
3.2 Initial Iteration with Large Residuals	15
3.3 Minimum-Distance Property	18
4. DISCUSSION	21
4.1 Basic Considerations	21
4.2 Additional Notes	24
5. SUMMARY AND CONCLUSIONS	28
APPENDIX A: NUMERICAL EXAMPLE	35
APPENDIX B: MODEL SURFACE AND MODEL PLANE	40
APPENDIX C: LINKS BETWEEN OBSERVATIONAL-SPACE, MODEL-PLANE, AND MODEL-SURFACE COORDINATE DIFFERENCES	44
APPENDIX D: MATRIX NOTATION	48
REFERENCES	51

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	



1. INTRODUCTION

The parametric adjustment model expresses each of the observables in terms of parameters, where the structure linking the two groups of variables is, in general, nonlinear. The number of observables is denoted by n and the number of parameters by u , where n must be greater than u for an adjustment to take place. The adjustment model is written as

$$L^a = F(X^a) ,$$

where L^a and X^a are the sets (column vectors) of adjusted observations and adjusted parameters, respectively. This study addresses the resolution of a nonlinear model through an isomorphic geometrical setup with tensor structure and notation. Such efforts date back to [Blaha, 1984], which treats a linear or linearized adjustment model. An initial analysis of a nonlinear model can be found in [Blaha, 1987].

In a standard adjustment approach, a nonlinear adjustment model is subject to the Taylor series expansion based on an initial set of parametric values, X^0 . The terms in the second and higher powers of the parametric corrections are neglected, resulting in the familiar (linearized) observation equations. In matrix notation, the latter are expressed by

$$V = AX + L ,$$

where A is the design matrix, $X = X^a - X^0$ is the column vector of parametric corrections, $V = L^a - L^b$ is the column vector of residuals, and $L = L^0 - L^b$ is the column vector of constant terms, with $L^0 = F(X^0)$ containing the values of observables consistent with the initial set of parameters and L^b containing the actual observations. The linearized model is subjected to the least-squares criterion

$$V^T P V = \text{minimum} ,$$

where P is the weight matrix of observations. This criterion leads to the formation of the familiar normal equations.

If the original adjustment model is nonlinear, the resolution of the linearized model does not yield the final answers. The process is usually repeated with new, updated parameters and the corresponding changes in A and L . However, the variance-covariance matrix of observations, Σ , as well as the

weight matrix P , adopted as $P = \Sigma^{-1}$, are independent of the parameters. Thus, the matrix of normal equations, $N = A^T P A$, changes only due to A , and the column vector representing the right-hand side of normal equations, $U = -A^T P L$, changes only due to A and L . The computation of the updated parametric values through a new X requires the formation and the inversion of a new N in each iteration, or a mathematically equivalent procedure. When X becomes sufficiently close to zero the iterative process is terminated. As its by-product, the latest matrix N^{-1} is adopted as the variance-covariance matrix of adjusted parameters.

The functional relationship between the observables and the parameters lends itself to a geometrical interpretation and treatment involving spaces and surfaces generalized to higher dimensions. In particular, the parametric adjustment model symbolized by

$$x^r = x^r(u^\alpha), \quad r = 1, 2, \dots, n, \quad \alpha = 1, 2, \dots, u$$

can be linked to the Gauss form of a surface in relation to the surrounding space, where x^r are the space coordinates and u^α are the surface coordinates. The Gauss form of a two-dimensional surface ($u=2$) embedded in a three-dimensional flat space ($n=3$) is described, together with two other forms, in Chapter 6 of [Hotine, 1969]. In [Blaha, 1984], both the n -dimensional "observational" space and the u -dimensional "model" surface were considered flat. The latter was thus in reality a hyperplane. Although the model surface is now intrinsically a curved space, the surrounding space will be seen to be again flat.

In general, a flat space can be described via Cartesian coordinates. In a Cartesian coordinate system, a given point is depicted by a set of coordinates x^r , $r=1, 2, \dots, n$, which can be interpreted as its position vector expressed by contravariant components, $\rho^r = x^r$. Considered as a tensor, and thus as a point function, ρ^r is associated with this given point (and not, for example, with the origin). Accordingly, if the space coordinates are Cartesian, the above model equation $x^r = x^r(u^\alpha)$ will express a family of position vectors associated with the model surface. Any of these position vectors could be freely parallel-transported to any location in space, e.g. to the surface point P described by $x_0^r = x^r(u_0^\alpha)$, and could eventually give rise to tensor equations there. Although derived in Cartesian coordinates, such equations would be valid in any coordinates applicable to a flat space.

In an important extension of the above discussion, we state that if

$$g_{sr} = \text{constant} ,$$

where g_{sr} is the space metric tensor, $x^r(u^\alpha)$ can again be interpreted as a family of position vectors associated with the model surface embedded in a flat space. Thus, although the space coordinate system cannot be general for such an interpretation to hold true, it is not required to be Cartesian. The position vectors can again be parallel-transported to the point P or any other location without changes in their components (contravariant as well as covariant), which has been referred to as "free" parallel transport. This stems from the fact that such changes are expressible in terms of the Christoffel symbols formed through partial derivatives of the space metric tensor.

Among all possible space coordinate systems, only those characterized by a constant metric tensor will be relevant to our development. Accordingly, $x^r(u^\alpha)$ will be interpreted as a family of position vectors associated with the model surface, any of which can be freely parallel-transported to a chosen location. The foregoing is meant to elucidate the isomorphism of adjustments and geometry. At the same time, it has led us to appreciate the qualitative difference between the two contexts. In the first context, the model equation $x^r = x^r(u^\alpha)$ provides restrictions on adjusted observations of a general kind, where all the quantities and relationships can be completely void of any geometrical meaning. In the second context, it provides restrictions on position vectors in space, which is a purely geometrical phenomenon. However, studying one scientific discipline in terms of another, even seemingly disparate, may lead to unforeseen benefits and a deeper understanding of both.

2. GEOMETRICAL SETUP

In denoting the n observables by x^r , $r=1,2,\dots,n$, and the u unknown parameters by u^α , $\alpha=1,2,\dots,u$, we can represent a nonlinear parametric adjustment model by

$$\begin{aligned} x^r = x^r(u^\alpha) = x_0^r &+ A_\alpha^r \Delta u^\alpha + (1/2) \Omega_{\alpha\beta}^r \Delta u^\alpha \Delta u^\beta \\ &+ (1/6) \Phi_{\alpha\beta\gamma}^r \Delta u^\alpha \Delta u^\beta \Delta u^\gamma + \dots, \end{aligned} \quad (1a)$$

$$\Delta u^\alpha = u^\alpha - u_0^\alpha, \quad (1b)$$

where u_0^α represents an initial set of parameters and $x_0^r = x^r(u_0^\alpha)$ represents the observables consistent with this set. Throughout this study, the lower-case Roman indices range from 1 to n , and the lower-case Greek indices range from 1 to u . Tensor symbolism implies the summation convention over the dummy (repeating) indices.

In the geometrical context, the first equality in (1a) represents the Gauss form of a u -dimensional surface embedded in an n -dimensional space. The surface is endowed with the coordinate system $\{u^\alpha\}$, $\alpha=1,2,\dots,u$ and is referred to as the model surface, and the space is endowed with the coordinate system $\{x^r\}$, $r=1,2,\dots,n$, and is referred to as the observational space. The second equality in (1a) is the Taylor series expansion of x^r from the "initial" point P lying in the model surface, whose model-surface coordinates are u_0^α and whose observational-space coordinates are x_0^r . The notation identifying the partial derivatives at P , such as $\partial x^r / \partial u^\alpha \equiv A_\alpha^r$, $\partial^2 x^r / \partial u^\alpha \partial u^\beta \equiv \Omega_{\alpha\beta}^r$, etc., is self-evident. The actual observations can be thought of as describing the point Q in the observational space, which, due to measuring errors, does not lie in the known model surface. The task at hand consists in determining, from the observed point Q , a model-surface point satisfying the least-squares criterion.

In the adjustment context, the variance-covariance and the weight matrices of observations depend on the quality of measurements. They are independent of the adjustment model, of the initial set of parameters, of the outcome of observations, etc. Thus, for a given observational mode they are constant. In the "traditional" identification of [Blaha, 1984], variance-covariance matrices correspond to associated metric tensors, and weight matrices correspond to metric tensors. Accordingly, we represent the variance-covariance matrix of

observations by the observational-space associated metric tensor g^{rs} , and the weight matrix of observations by the observational-space metric tensor g_{sr} , and state that both tensors are independent of the form of the model surface, of the initial point P, of the observed point Q, etc., leading to the simplification

$$g_{sr} = \text{constant} . \quad (2)$$

One could also attribute the tensors g^{rs} and g_{sr} to the point Q and state that the geometrical setup must account for Q located anywhere in the observational space. In turn, (2) implies that the observational space must be flat.

If the set δx^r denotes the coordinate differences between the observed point Q and the desired model-surface point denoted \bar{P} , it corresponds to the negative residuals, and the least-squares criterion corresponds to

$$\delta s^2 = \delta x^s g_{sr} \delta x^r = \text{minimum} . \quad (3)$$

Since (2) allows us to make appeal to Cartesian coordinates, we can consider the quantities such as δx^r , Δx^r , etc., whether infinitesimal or finite, to be sets of coordinate differences as well as contravariant vectors in the observational space, where they can be freely parallel-transported to any location. Moreover, the quadratic form (3) represents the square of the distance between Q and \bar{P} . Therefore, the desired "least-squares" point \bar{P} must be the foot-point of the straight line dropped orthogonally from Q onto the model surface. This property is schematically illustrated in Fig. 1, where the vector $\bar{P}Q$, i.e., the geometrical object symbolized by δx , is drawn perpendicular to the model surface at \bar{P} . We note that if any other adjustment condition were used in lieu of the least-squares criterion, the minimum-distance property (3) would not exist and the geometric-tensorial treatment of the adjustment theory would probably be much more complex if not impossible.

The model-surface coordinates of the least-squares point \bar{P} , denoted u^α , correspond to the adjusted parameters, and its observational-space coordinates, denoted x^r , correspond to the adjusted observations. (Although this designation is more restrictive than the notation u^α and x^r used in 1a,b, it need not cause any confusion.) Furthermore, the surface associated metric tensor at \bar{P} corresponds to the variance-covariance matrix of adjusted parameters, and the "necessary" version of the space associated metric tensor at \bar{P} corresponds to the singular variance-covariance matrix of adjusted observations. Such

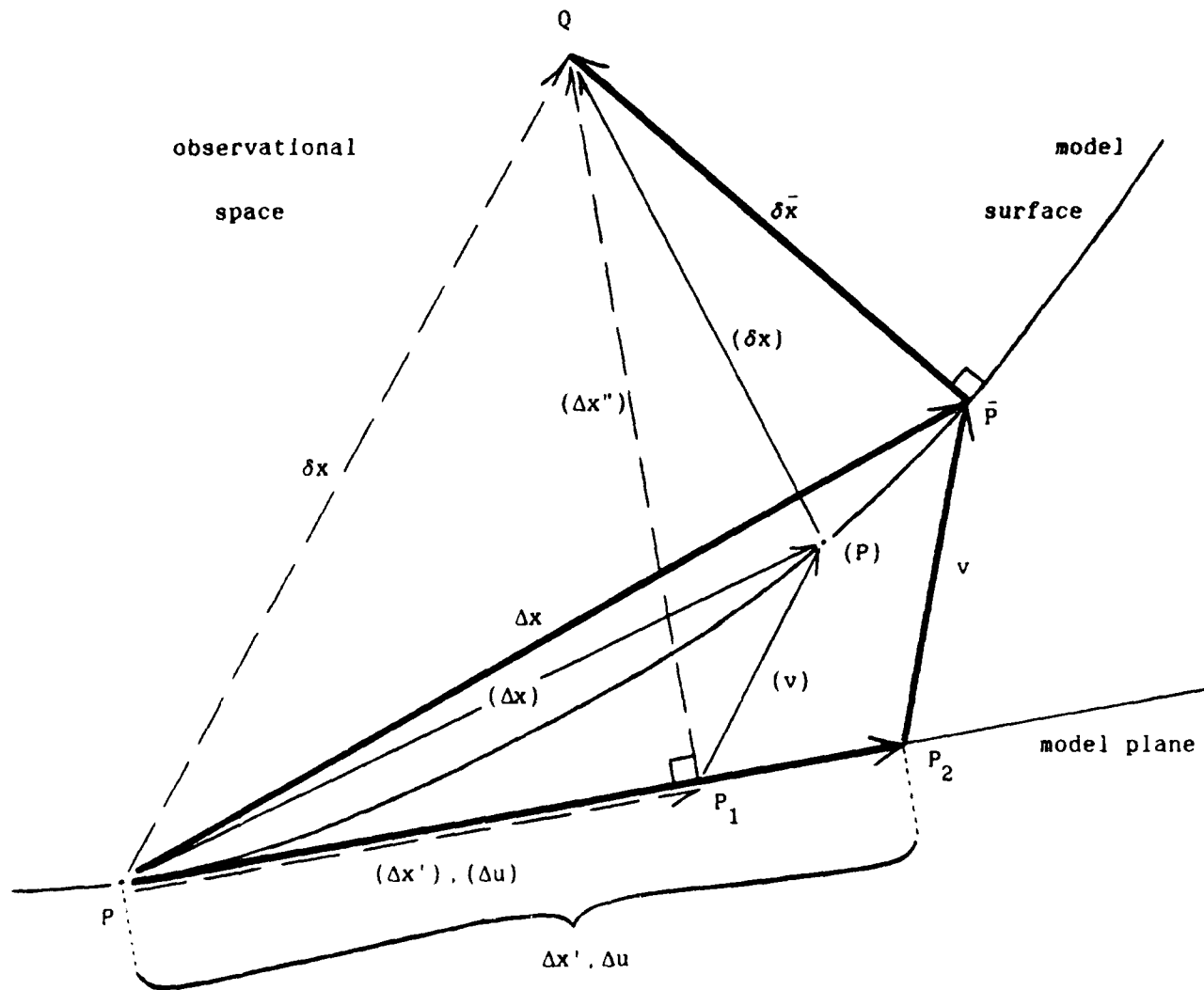


Fig. 1

Symbolic representation of a nonlinear parametric least-squares adjustment. A one-dimensional model surface is embedded in a two-dimensional observational space; the parentheses pertain to the solution of the model linearized at P.

correspondences, as well as the definition of "necessary" tensors, were presented in Sections 2.2 and 4.2 of [Blaha, 1984]. In the same vein, the surface metric tensor at \bar{P} corresponds to the weight matrix of adjusted parameters, and the "necessary" version of the space metric tensor at \bar{P} corresponds to the singular weight matrix of adjusted observations.

These notions confirm the variance-covariance propagation law and introduce a "weight propagation law", and they can be extended to functions of adjusted parameters as well as to functions of adjusted observations. In [Blaha, 1984], however, they were demonstrated only for linear models. In nonlinear models the above tensors, expressible by means of orthonormal vectors tangent to the model surface at \bar{P} , could be said to describe "tangential", or first-order, properties of least-squares estimates.

3. SOLUTION

3.1 General Development

The adjustment theory assumes that the known values of the initial set of parameters, represented here by u_0^α , are close to the final values represented by u^α . Furthermore, the values of the residuals are also assumed to be small. For the needs of the current development, we stipulate that quantities will be considered "small", or first-order, if they behave like differential quantities, in the sense that a product of two first-order quantities will result in a quantity smaller by an order of magnitude. The latter will be referred to as a second-order quantity. In the geometrical context the assumption about u_0^α is maintained, implying that the model-surface coordinate differences between \bar{P} and P , in (1b) symbolized by Δu^α , are small. The same assumption applies also for the observational-space coordinate differences between these two points,

$$\Delta x^r = x^r - x_0^r. \quad (4)$$

However, the values in $\delta \bar{x}^r \equiv x_Q^r - x^r$, corresponding to the negative residuals, are subject to no such restriction. This is the first distinction between the standard and the geometrical approaches to nonlinear models.

The second distinction, which is the cornerstone of our geometrical approach, is a consequence of (3). Before drawing conclusions from this minimum-distance criterion, we review the notion of the design tensor, corresponding in adjustment calculus to the design matrix. In [Blaha, 1984] and in Section 2.2 of [Blaha, 1987], this tensor was identified with the set of partial derivatives $\partial x^r / \partial u^\alpha$ at the initial point P , and was presented in a vectorized form. Thus, A_α^r in (1a) is the design tensor at P , which can be expressed by means of a set of orthonormal vectors spanning the hyperplane tangent to the model surface at P :

$$A_\alpha^r \equiv \partial x^r / \partial u^\alpha = \ell^r \ell_\alpha + j^r j_\alpha + \dots \quad (5)$$

The above hyperplane is referred to as the "model plane", and in Fig.1 it is symbolically represented by a straight line. The notation ℓ^r, j^r, \dots identifies the orthonormal vectors at P spanning the model plane by their observational-space contravariant components, while the notation $\ell_\alpha, j_\alpha, \dots$ identifies them by their model-surface covariant components. In general, when

identifying individual vectors as geometrical objects, we remove the indices. The double identification, such as $\Delta x'$ and Δu written side-by-side in Fig. 1, depicts one and the same vector expressed in two coordinate systems. The notions of model surface and model plane are further discussed in Appendix B, which also describes in detail the role of the design tensor in either manifold.

The design tensor (5) serves in formulating the model-surface metric tensor at the initial point P as

$$a_{\beta\alpha} = \ell_{\beta} \ell_{\alpha} + j_{\beta} j_{\alpha} + \dots = A_{\beta}^s g_{sr} A_{\alpha}^r. \quad (6)$$

Relations similar to (5) and (6) appear in Chapter 6 of [Hotine, 1969], where they are derived for $n=3$ and $u=2$. Equation (5) provided an early impetus in treating the adjustment theory via geometry. Its motivating power stems from the fact that its left-hand side describes the algebraic structure of the design matrix in the parametric adjustment model, and its right-hand side gives this structure a clearcut geometrical meaning.

In analogy to A_{α}^r , the design tensor at \bar{P} can be formulated as

$$\bar{A}_{\alpha}^r = \bar{\ell}^r \bar{\ell}_{\alpha} + \bar{j}^r \bar{j}_{\alpha} + \dots$$

where the overbars identify a set of orthonormal vectors spanning the hyperplane tangent to the model surface at \bar{P} . Due to the orthogonality of $\delta \bar{x}$ and the model surface as dictated by (3), at \bar{P} we must have

$$\bar{A}_{\beta}^s g_{sr} \delta \bar{x}^r = 0. \quad (7)$$

Since \bar{A}_{β}^s is a mixed space-surface tensor at \bar{P} , and $\delta \bar{x}^r$ is a contravariant space vector at \bar{P} , the above is a tensor equation at \bar{P} . This equation appeared in a similar form in Section 2.3 of [Blaha, 1987]. The orthogonality condition (7) embodies the above-mentioned second distinction, in that it is a precise relation whose active exploitation does not have an equivalent in the standard adjustment approach. As has been indicated in the Introduction, the latter proceeds by iterations, where a set of values approximating Δu^{α} is refined at each step. When the changes are sufficiently small, the iterative process terminates and (7) is satisfied as a by-product. However, upon seeking to fulfill this condition at each individual step, the standard approach can be modified and the iterative process can be significantly shortened.

Since, in the geometrical context, P is assumed to be close to \bar{P} , a good approximation to Δu^α is offered by (Δu^α) , where the parenthesis notation represents the solution of the linearized model described by (1a,b) with $\Omega_{\alpha\beta}^r = 0$, $\Phi_{\alpha\beta\gamma}^r = 0$, This approximation is, in fact, the initial iteration of the standard approach. We note that if the values in δx^r are so large as to approach finite quantities, a better approximation to Δu^α can be obtained in a different, more general form presented in the next section. In the current discussion the linearized model corresponds to the situation where the model plane replaces the model surface. We retain $\{u^\alpha\}$ as the model-plane coordinate system, and observe that (6) is the plane's constant metric tensor. In analogy to a statement made earlier in conjunction with the constant space metric tensor, we can consider the quantities such as (Δu^α) , Δu^α , etc., whether infinitesimal or finite, to be sets of model-plane coordinate differences as well as contravariant vectors in the model plane, where, if need be, they could be freely parallel-transported to any location.

The model-plane covariant components of the orthogonal projection of δx^r onto the model plane are obtained with the aid of the design tensor (5):

$$(\Delta u_\beta) = A_\beta^s g_{sr} \delta x^r, \quad (8)$$

where the contravariant vector δx^r , known a priori, can be decomposed as

$$\delta x^r \equiv x_Q^r - x_O^r = (\Delta x'^r) + (\Delta x''^r).$$

In terms of Fig. 1, the double identification $(\Delta x')$ and (Δu) designates one and the same vector PP_1 . Based on (8), we have

$$(\Delta u^\alpha) = a^{\alpha\beta} (\Delta u_\beta), \quad (9)$$

where $a^{\alpha\beta}$ is the model-surface associated metric tensor at P , or the constant model-plane associated metric tensor, obtained from $a_{\beta\alpha}$ according to

$$a^{\alpha\beta} a_{\beta\gamma} = \delta_\gamma^\alpha.$$

If (Δu^α) is considered to be a set of model-surface rather than model-plane coordinate differences, it gives rise to a surface point which, in Fig. 1, is denoted (P) . In consulting the figure, we note that in the observational space

this point is arrived at by means of the vector (Δx) , where

$$(\Delta x^r) = (\Delta x'^r) + (v^r) .$$

Topics similar to the above are discussed in Appendix C.

If the adjustment model were linear, the model surface would reduce to the model plane and (Δu^α) would represent the final parametric corrections. The design tensor (5) would serve to express

$$(\Delta x'^r) = A_\alpha^r (\Delta u^\alpha) ,$$

which would yield the adjusted observations as $x_0^r + (\Delta x'^r)$. The negative residuals would be found from

$$(\Delta x''^r) = \delta x^r - (\Delta x'^r) .$$

The variance-covariance matrix of adjusted parameters would correspond to $a^{\alpha\beta}$, while the variance-covariance matrix of adjusted observations would correspond to the necessary associated metric tensor:

$$g'^{rs} = l^r l^s + j^r j^s + \dots = A_\alpha^r a^{\alpha\beta} A_\beta^s . \quad (10a)$$

The weight matrix of adjusted parameters would correspond to $a_{\beta\alpha}$ and the weight matrix of adjusted observations would correspond to the necessary metric tensor:

$$g'_{sr} = g_{s1} g'^{1j} g_{jr} . \quad (10b)$$

Furthermore, the variance-covariance and the weight matrices of residuals would correspond respectively to

$$g''^{rs} = g^{rs} - g'^{rs} , \quad g''_{sr} = g_{sr} - g'_{sr} . \quad (11a,b)$$

The above relations were presented in [Blaha, 1984].

In the current nonlinear model, (Δu^α) represents the first approximation to Δu^α . We remark that in a two- or higher-dimensional model plane the vectors (Δu) and Δu are not collinear in general. The set Δu^α , when considered in the role of model-surface coordinate differences, gives rise to the desired least-squares point \bar{P} (see Fig. 1). In the observational space this point is obtained by means of the vector Δx , where

$$\Delta x^r = \Delta x'^r + v^r . \quad (12a)$$

From (1a,b) and (4), we express the right-hand side of (12a) by

$$\Delta x'^r = A_\alpha^r \Delta u^\alpha, \quad (12b)$$

$$v^r = (1/2) \Omega_{\alpha\beta}^r \Delta u^\alpha \Delta u^\beta + (1/6) \Phi_{\alpha\beta\gamma}^r \Delta u^\alpha \Delta u^\beta \Delta u^\gamma + \dots \quad (12c)$$

This outcome is also discussed in Appendix C.

Whether the least-squares point \bar{P} should be described by the model-surface coordinates or by the observational-space coordinates via (12a-c), the values Δu^α are unknown and unobtainable by closed formulas such as (8) and (9) applicable to the linear model. In view of formulating a useful relationship for Δu^α , we first observe that

$$\bar{A}_\alpha^r = A_\alpha^r + \Delta A_\alpha^r, \quad (13a)$$

$$\Delta A_\alpha^r = \Omega_{\alpha\beta}^r \Delta u^\beta + (1/2) \Phi_{\alpha\beta\gamma}^r \Delta u^\beta \Delta u^\gamma + \dots, \quad (13b)$$

where (13b) stems from the fact that at P we have $\partial A_\alpha^r / \partial u^\beta = \Omega_{\alpha\beta}^r$, $\partial^2 A_\alpha^r / \partial u^\beta \partial u^\gamma = \Phi_{\alpha\beta\gamma}^r$, etc. Equations (13a,b) express the design tensor at \bar{P} from its counterpart at P. Further, Fig. 1 shows that $\delta \bar{x}^r$ can be expressed as

$$\delta \bar{x}^r = \delta x^r - \Delta x^r, \quad (14)$$

where δx^r is known but Δx^r , determined via Δu^α in (12a-c), is unknown. This is a tensor equation at the least-squares point \bar{P} , to which δx^r and Δx^r can be freely parallel-transported from the initial point P. The tensors \bar{A}_α^r and $\delta \bar{x}^r$ at \bar{P} are needed in view of (7).

The approximation set (Δu^α) will be corrected by $\Delta(\Delta u^\alpha)$ to yield the desired set Δu^α :

$$\Delta u^\alpha = (\Delta u^\alpha) + \Delta(\Delta u^\alpha). \quad (15)$$

The unknown set $\Delta(\Delta u^\alpha)$ will be determined through the precise relation (7). We stipulate that the level of accuracy in (7) should be of the third order, in the sense that the neglected terms can only be of the fourth and higher orders. By a hypothesis common to both the standard and the geometrical approaches, the set Δu^α contains small quantities and the set (Δu^α) represents the main contribution of Δu^α . Thus, (Δu^α) and Δu^α are assumed to contain first-order quantities, while the set $\Delta(\Delta u^\alpha)$ is assumed to contain second-order quantities. Finite quantities contracted with $(\Delta u^\alpha)(\Delta u^\beta)$ are considered to be of the second order.

similar to finite quantities contracted with $\Delta(\Delta u^\alpha)$. Fourth-order quantities to be neglected are of three kinds: $o[(\Delta u^\alpha)(\Delta u^\beta)(\Delta u^\gamma)(\Delta u^\delta)]$, $o[(\Delta u^\alpha)(\Delta u^\beta) \Delta(\Delta u^\gamma)]$, and $o[\Delta(\Delta u^\alpha) \Delta(\Delta u^\beta)]$. Since the last kind involves quadratic terms in $\Delta(\Delta u^\alpha)$, the reason for carrying out the solution to the third order becomes clearer.

In considering the known quantities (Δu^α) , similar to (13a,b) we have

$$(A_\alpha^r) = A_\alpha^r + \Omega_{\alpha\beta}^r (\Delta u^\beta) + (1/2) \Phi_{\alpha\beta\gamma}^r (\Delta u^\beta) (\Delta u^\gamma) + \dots \quad (16)$$

We now form \bar{A}_α^r based on the outcome for (A_α^r) :

$$\bar{A}_\alpha^r = (A_\alpha^r) + \Delta(A_\alpha^r) \quad (17a)$$

The set $\Delta(A_\alpha^r)$ can be obtained as \bar{A}_α^r in (13a,b) with Δu^α substituted for by the right-hand side of (15), minus (A_α^r) from (16). This yields

$$\Delta(A_\alpha^r) = \Omega_{\alpha\beta}^r \Delta(\Delta u^\beta) + \Phi_{\alpha\beta\gamma}^r (\Delta u^\gamma) \Delta(\Delta u^\beta) + \dots \quad (17b)$$

where use has been made of the symmetry of $\Phi_{\alpha\beta\gamma}^r$ in α , β , and γ through

$$\Phi_{\alpha\beta\gamma}^r \Delta(\Delta u^\beta) (\Delta u^\gamma) + \Phi_{\alpha\beta\gamma}^r (\Delta u^\beta) \Delta(\Delta u^\gamma) = 2 \Phi_{\alpha\beta\gamma}^r (\Delta u^\gamma) \Delta(\Delta u^\beta) .$$

The dots in (17b) represent $o[\Delta(\Delta u^\gamma) \Delta(\Delta u^\beta)] + o[(\Delta u^\delta)(\Delta u^\gamma) \Delta(\Delta u^\beta)] + \dots$, which are respectively fourth-, fourth-, and higher-order quantities.

Furthermore, similar to (12a-c) we write

$$\begin{aligned} (\Delta x^r) &= A_\alpha^r (\Delta u^\alpha) + (1/2) \Omega_{\alpha\beta}^r (\Delta u^\alpha) (\Delta u^\beta) \\ &\quad + (1/6) \Phi_{\alpha\beta\gamma}^r (\Delta u^\alpha) (\Delta u^\beta) (\Delta u^\gamma) + \dots \end{aligned} \quad (18)$$

In analogy to the above, we form Δx^r based on the outcome for (Δx^r) :

$$\Delta x^r = (\Delta x^r) + \Delta(\Delta x^r) \quad (19a)$$

The set $\Delta(\Delta x^r)$ can be obtained as Δx^r in (12a-c) with Δu^α given by (15), minus (Δx^r) given by (18). This yields

$$\Delta(\Delta x^r) = A_\alpha^r \Delta(\Delta u^\alpha) + \Omega_{\alpha\beta}^r (\Delta u^\beta) \Delta(\Delta u^\alpha) + \dots \quad (19b)$$

where we have employed the symmetry of $\Omega_{\alpha\beta}^r$ in α and β through the identity

$$\Omega_{\alpha\beta}^r \Delta(\Delta u^\alpha) (\Delta u^\beta) + \Omega_{\alpha\beta}^r (\Delta u^\alpha) \Delta(\Delta u^\beta) = 2 \Omega_{\alpha\beta}^r (\Delta u^\beta) \Delta(\Delta u^\alpha) .$$

However, due to (16), equation (19b) simplifies to

$$\Delta(\Delta x^r) = (A_\alpha^r) \Delta(\Delta u^\alpha) + \dots, \quad (19c)$$

where the dots represent $o[\Delta(\Delta u^\beta) \Delta(\Delta u^\alpha)] + o[(\Delta u^\delta)(\Delta u^\gamma)(\Delta u^\beta) \Delta(\Delta u^\alpha)] + \dots$, which are respectively fourth-, fifth-, and higher-order quantities.

Finally, due to (14) and (19a), we can express δx^r as

$$\delta x^r = (\delta x^r) - \Delta(\Delta x^r), \quad (20a)$$

where

$$(\delta x^r) = \delta x^r - (\Delta x^r). \quad (20b)$$

In terms of (17a) and (20a), the precise relation (7) reads

$$[(A_\beta^s) + \Delta(A_\beta^s)] g_{sr} [(\delta x^r) - \Delta(\Delta x^r)] = 0,$$

or

$$\begin{aligned} & (A_\beta^s) g_{sr} (\delta x^r) + \Delta(A_\beta^s) g_{sr} (\delta x^r) \\ & - (A_\beta^s) g_{sr} \Delta(\Delta x^r) - \Delta(A_\beta^s) g_{sr} \Delta(\Delta x^r) = 0. \end{aligned} \quad (21)$$

The second, third, and fourth terms on the left-hand side of (21) are denoted respectively as a, b, and c, and will now be evaluated.

In considering (17b) we obtain

$$a = \Omega_{\beta\gamma}^s \Delta(\Delta u^\gamma) g_{sr} (\delta x^r) + \Phi_{\beta\gamma\delta}^s (\Delta u^\delta) \Delta(\Delta u^\gamma) g_{sr} (\delta x^r) + \dots$$

In analogy to the statement that followed (17b), the dots represent fourth- and higher-order quantities. However, (δx^r) is excluded from influencing this order. Further, in consulting (19c) we write

$$b = (A_\beta^s) g_{sr} (A_\alpha^r) \Delta(\Delta u^\alpha) + \dots$$

In analogy to the statement that followed (19c), the dots again represent fourth- and higher-order quantities. Finally, in consulting both (17b) and (19c) we conclude that

$$c = o[\Delta(\Delta u^\gamma) \Delta(\Delta u^\alpha)] + \dots$$

which is composed entirely of fourth- and higher-order quantities.

Upon collecting the results, (21) yields

$$\begin{aligned} (A_{\beta}^s) g_{sr} (\delta x^r) = & (A_{\beta}^s) g_{sr} (A_{\alpha}^r) \Delta(\Delta u^{\alpha}) - \Omega_{\beta\alpha}^s \Delta(\Delta u^{\alpha}) g_{sr} (\delta x^r) \\ & - \Phi_{\beta\alpha\gamma}^s (\Delta u^{\gamma}) \Delta(\Delta u^{\alpha}) g_{sr} (\delta x^r) + \dots, \end{aligned} \quad (22)$$

where the dots represent fourth- and higher-order quantities regardless of (δx^r) . Equation (22) is equivalent to the precise relation (7). If the terms symbolized by dots are omitted, the equivalence is valid to the third order. In this case, the desired set $\Delta(\Delta u^{\alpha})$ is obtained by solving the linear system

$$[(a_{\beta\alpha}) - (\delta x_s) \Omega_{\beta\alpha}^s - (\delta x_s) \Phi_{\beta\alpha\gamma}^s (\Delta u^{\gamma})] \Delta(\Delta u^{\alpha}) = (A_{\beta}^s) (\delta x_s), \quad (23a)$$

where

$$(a_{\beta\alpha}) = (A_{\beta}^s) g_{sr} (A_{\alpha}^r), \quad (\delta x_s) = g_{sr} (\delta x^r). \quad (23b,c)$$

All the quantities in parentheses are formed with (Δu^{α}) . The relation (23a) could be called "modified normal equations", and the matrix form of the quantity inside the brackets could be called "matrix of modified normal equations". The second- and the third-order partial derivatives grouped respectively in $\Omega_{\beta\alpha}^s$ and $\Phi_{\beta\alpha\gamma}^s$ are evaluated at P and can thus be stored and treated as constant should (23a) be re-applied. Their "once and for all" evaluation is useful since in complex models their re-computation at an updated point would represent a serious drawback. We remark that if (δx^r) is small, in the sense it contains first-order quantities, the third term inside the brackets of (23a) can be left out since it would give rise to fourth-order quantities.

3.2 Initial Iteration with Large Residuals

The development in the previous section was based on the set (Δu^{α}) as obtained in the initial iteration of the standard adjustment approach. In [Blaha, 1987], such a common iteration was called "zero-th iteration". This concept is abandoned since the two approaches can now be considered distinct already at the initial stage. In particular, we develop a general case of the geometrical approach, where already the first, or initial, iteration is based on the exact relation (7) valid at the least-squares point \bar{P} .

In consulting the previous section, we present the barred tensors in (7):

$$\bar{A}_{\beta}^s = A_{\beta}^s + \Omega_{\beta\gamma}^s \Delta u^{\gamma} + (1/2) \Phi_{\beta\gamma\delta}^s \Delta u^{\gamma} \Delta u^{\delta} + \dots, \quad (24a)$$

$$\delta \bar{x}^{\Gamma} = \delta x^{\Gamma} - A_{\alpha}^{\Gamma} \Delta u^{\alpha} - (1/2) \Omega_{\alpha\epsilon}^{\Gamma} \Delta u^{\alpha} \Delta u^{\epsilon} - \dots \quad (24b)$$

To avoid expressions nonlinear in Δu^{α} , we introduce approximations in (7) whereby all terms containing the second and higher powers in Δu^{α} will be left out, i.e., $o(\Delta u^{\alpha} \Delta u^{\beta})$, $o(\Delta u^{\alpha} \Delta u^{\beta} \Delta u^{\gamma})$, etc., will be neglected. This means that the third and further terms on the right-hand sides of (24a,b) will be neglected, as well as the term $-\Omega_{\beta\gamma}^s \Delta u^{\gamma} g_{sr} A_{\alpha}^r \Delta u^{\alpha}$ participating in (7). Equation (7) is thus written as

$$A_{\beta}^s g_{sr} \delta x^{\Gamma} - A_{\beta}^s g_{sr} A_{\alpha}^{\Gamma} \Delta u^{\alpha} + \Omega_{\beta\gamma}^s \Delta u^{\gamma} g_{sr} \delta x^{\Gamma} + \dots = 0, \quad (25)$$

where the dots represent the terms which would contain the second and higher powers of Δu^{α} , and where no assumptions have been made yet about the size of the known contravariant vector δx^{Γ} . We add that the values constituting A_{α}^{Γ} , $\Omega_{\alpha\beta}^{\Gamma}$, and g_{sr} are assumed finite and of mutually comparable magnitudes.

Upon the use of the symbols $a_{\beta\alpha}$ and δx_s , where

$$a_{\beta\alpha} = A_{\beta}^s g_{sr} A_{\alpha}^{\Gamma}, \quad \delta x_s = g_{sr} \delta x^{\Gamma}, \quad (26a,b)$$

equation (25) becomes

$$[a_{\beta\alpha} - \delta x_s \Omega_{\beta\alpha}^s + \dots] \Delta u^{\alpha} = A_{\beta}^s \delta x_s. \quad (27)$$

The first geometrical iteration is characterized by the neglect of the second and higher powers of Δu^{α} in (27), in which case the solution of Δu^{α} is denoted temporarily as $\{\Delta u^{\alpha}\}$:

$$[a_{\beta\alpha} - \delta x_s \Omega_{\beta\alpha}^s] \{\Delta u^{\alpha}\} = A_{\beta}^s \delta x_s. \quad (27')$$

In terms of Fig. 1, this solution would lead to a model-plane point lying somewhere between P_1 and P_2 , and to a model-surface point lying somewhere between (P) and \bar{P} . If we further neglect $-\delta x_s \Omega_{\beta\alpha}^s \Delta u^{\alpha}$ and denote the corresponding solution by (Δu^{α}) , equation (27) becomes

$$a_{\beta\alpha} (\Delta u^{\alpha}) = A_{\beta}^s \delta x_s, \quad (27'')$$

which symbolizes the normal equations in the standard adjustment approach.

In the standard approach, the initial values of parameters, u_0^α , are assumed to be close to the final values, u^α . In our current notation and terminology, this assumption, which we shall retain, is equivalent to stating that Δu^α contains first-order quantities. Since only the second- and higher-order quantities are ignored in (27'), $\{\Delta u^\alpha\}$ in the first geometrical iteration will recover the main contribution of Δu^α . The errors in $\{\Delta u^\alpha\}$ will thus have the magnitude of second-order quantities, consistent with the property previously attributed to the initial standard iteration.

As the comparison of (27'') with (27') reveals, this property of the initial standard iteration is indeed valid, provided the term $-\delta x_s \Omega_{\beta\alpha}^s \Delta u^\alpha$ neglected in the standard approach is at the level of second-order quantities, i.e., at the level of the terms already neglected by (27'). In general, this is true if δx^r contains first-order quantities, which is another assumption inherent in the standard approach. Accordingly, under the two standard assumptions, requiring that both Δu^α and δx^r contain first-order quantities, the solution of (27'') is valid to the first order, similar in this respect to the solution of (27'). The errors in both (Δu^α) and $\{\Delta u^\alpha\}$ are then second-order quantities.

If the second assumption above is not satisfied and δx^r contains finite quantities, (27') will still hold true to the first order whereas (27'') will not. Although the quantities in (Δu^α) will be of the first order, their errors will in general attain the first order as well, due to the neglect of the first-order quantities $-\delta x_s \Omega_{\beta\alpha}^s \Delta u^\alpha$. This case is mainly of theoretical interest, since, in practice, it corresponds to excessive observational errors which should be eliminated beforehand. However, we might also consider the instances where δx^r belongs to a category "in-between" and contains quantities that are finite but small, without being small enough to warrant the attribute "first order". Here the solution of (27') is better than the solution of (27''), but not by an order of magnitude. We conclude that in general the solution of (27') approximates Δu^α as well as, or better than, the solution of (27''). Better results are obtained when the values in δx^r have the magnitude greater than that characterizing the first order.

The quantities forming δx^r and their link to (27') and (27'') can be further elaborated on as follows. Since Δu^α is of the first order by assumption, and since the same applies to $\{\Delta u^\alpha\}$ and (Δu^α) , equations (27') and (27'') reveal that

the tensor $A_{\beta}^s \delta x_s$, which figures on their right-hand sides, is also of the first order. In fact, (27'') allows this tensor to be written as

$$A_{\beta}^s g_{sr} \delta x^r = (\Delta u_{\beta}) .$$

The above expression identifies the projection of the vector δx onto the model plane, as previously depicted in Fig. 1. Since (Δu_{β}) is of the first order, it follows that if the magnitude of the vector δx is significantly greater than that characterizing the first order, this vector must be fairly nearly orthogonal to the model surface. This, in turn, imposes restrictions on the observational errors.

In summary, if δx^r is of the first order, which corresponds to the presence of small observational errors as encountered in practice, the first geometrical iteration (27') has the same level of accuracy as the first standard iteration (27''). This category includes $\delta x^r = 0$, which, in turn, includes the case of error-free observations. If the characteristics of δx^r fall into the "in-between" category, which corresponds to the presence of large observational errors, (27') has a higher level of accuracy than (27''), although not by an order of magnitude. Here the errors are restricted in the sense that δx must be fairly nearly orthogonal to the model surface. Finally, if δx^r contains finite values, which corresponds to the presence of excessively large observational errors (blunders) and is thus mainly of theoretical interest, (27') is still valid to the first order, whereas (27'') may give completely false results. In this case, the errors are restricted in the sense that δx must be very nearly orthogonal to the model surface. Since (27') gives results that are as good as those of (27'') in most cases, and better in the remaining cases, the use of this equation as the first geometrical iteration is fully justified.

3.3 Minimum-Distance Property

For the sake of theory, we express δs^{-2} (the square of the distance between \bar{P} and Q) by means of the Taylor series based on δs^2 (the square of the distance between P and Q). The latter is a known constant, namely

$$\delta s^2 = \delta x^s g_{sr} \delta x^r ,$$

while the former is a function of Δu^α , expressed as

$$\delta s^{-2} = \delta \bar{x}^s g_{sr} \delta \bar{x}^r,$$

where $\delta \bar{x}^r$ is given e.g. in (24b). In using the symmetry of g_{sr} and $\Omega_{\alpha\beta}^r$ in the lower indices, we deduce that

$$\delta s^{-2} = \delta s^2 + q_\alpha \Delta u^\alpha + (1/2) H_{\alpha\beta} \Delta u^\alpha \Delta u^\beta + \dots,$$

where

$$q_\alpha = \partial \delta s^{-2} / \partial u^\alpha = -2 A_\alpha^s g_{sr} \delta \bar{x}^r,$$

$$H_{\alpha\beta} = \partial^2 \delta s^{-2} / \partial u^\alpha \partial u^\beta = 2 (A_\beta^s g_{sr} A_\alpha^r - \delta \bar{x}^s g_{sr} \Omega_{\alpha\beta}^r), \dots$$

The second-order partial derivatives grouped in $H_{\alpha\beta}$ are assumed continuous, at least in the neighborhood of the point \bar{P} .

A similar expression can be written for a model-surface point in the neighborhood of \bar{P} , whose square of the distance from Q , denoted $\delta \bar{s}^2$, can be expressed in terms of $\delta \bar{s}^2$:

$$\delta \bar{s}^2 = \delta s^2 + \bar{q}_\alpha \Delta \bar{u}^\alpha + (1/2) \bar{H}_{\alpha\beta} \Delta \bar{u}^\alpha \Delta \bar{u}^\beta + \dots$$

Here $\Delta \bar{u}^\alpha$ is the difference, in model-surface coordinates, between the point in question and the point \bar{P} . The necessary and sufficient conditions for $\delta \bar{s}^2$ to have a local minimum at \bar{P} are

$$\bar{q}_\alpha = -2 \bar{A}_\alpha^s g_{sr} \delta \bar{x}^r = 0, \quad (28)$$

and

$$[\bar{H}_{\alpha\beta}] = \text{positive-definite matrix}; \quad (29)$$

the brackets indicate the matrix equivalent of $\bar{H}_{\alpha\beta}$, where

$$\bar{H}_{\alpha\beta} = 2 (\bar{A}_\beta^s g_{sr} \bar{A}_\alpha^r - \delta \bar{x}^s g_{sr} \bar{\Omega}_{\alpha\beta}^r).$$

The latter equation can be written as

$$(1/2) \bar{H}_{\alpha\beta} = \bar{a}_{\beta\alpha} - \delta \bar{x}_s \bar{\Omega}_{\beta\alpha}^s, \quad (30)$$

with

$$\bar{a}_{\beta\alpha} = \bar{A}_\beta^s g_{sr} \bar{A}_\alpha^r, \quad \delta \bar{x}_s = g_{sr} \delta \bar{x}^r. \quad (31a, b)$$

We first notice that (28) is satisfied by virtue of (7). Further, since

$$\bar{\Omega}_{\beta\alpha}^s = \Omega_{\beta\alpha}^s + \Phi_{\beta\alpha\gamma}^s \Delta u^\gamma + \dots, \quad (32)$$

equation (30) becomes

$$(1/2) \bar{H}_{\alpha\beta} = \bar{a}_{\beta\alpha} - \delta\bar{x}_s \Omega_{\beta\alpha}^s - \delta\bar{x}_s \Phi_{\beta\alpha\gamma}^s \Delta u^\gamma - \dots. \quad (33)$$

According to (29), the above quantity, symmetric in the indices α and β , should represent a positive-definite matrix. If we now compare the right-hand side of (33) with the quantity inside the brackets of (23a) in its final update, i.e., in the last iteration where $(a_{\beta\alpha})$, (δx_s) , and (Δu^γ) become essentially $\bar{a}_{\beta\alpha}$, $\delta\bar{x}_s$, and Δu^γ , we observe that within the required accuracy the two forms are identical. It is thus confirmed that the algorithm represented by (23a) leads to a minimum provided the updated matrix of modified normal equations is positive-definite, and vice-versa. Since the existence of a minimum is assumed throughout, and since P is assumed to be sufficiently close to \bar{P} , the above outcome implies that the matrix of modified normal equations will be positive-definite in all iterations, with or without the contribution of the third-order partial derivatives. This, in turn, implies that it can be inverted by the Choleski algorithm, as is illustrated in the example presented in Appendix A.

4. DISCUSSION

4.1 Basic Considerations

The leading subject of this discussion is equation (23a), applied after the initial iteration. We begin by relating (δx^r) to $\delta \bar{x}^r$, where, in the adjustment context, the latter represents the negative residuals. From (20a) we write

$$(\delta x^r) - \delta \bar{x}^r = \Delta(\Delta x^r) ,$$

which, upon using (19c), becomes

$$(\delta x^r) = \delta \bar{x}^r + (A_\alpha^r) \Delta(\Delta u^\alpha) + \dots \quad (34)$$

Here the dots again represent fourth- and higher-order quantities. Thus, (δx^r) differs from $\delta \bar{x}^r$ by second- and higher-order quantities.

We distinguish three basic cases with regard to the magnitude of the quantities comprising the set $\delta \bar{x}^r$:

- 1) If $\delta \bar{x}^r$ is of the second order or smaller, (δx^r) is in general of the second order and the two sets (of very small values) may be quite different. This case includes $\delta \bar{x}^r = 0$, where the point Q lies on the model surface as would happen, for example, with error-free observations.
- 2) If $\delta \bar{x}^r$ is of the first order, so is (δx^r) , and the two sets are similar in values.
- 3) If $\delta \bar{x}^r$ is large (finite), so is (δx^r) , and the two sets are nearly equal in values.

In the first case, (δx_s) is of the second order and (23a) simplifies to

$$(a_{\beta\alpha}) \Delta(\Delta u^\alpha) = (A_\beta^s) (\delta x_s) . \quad (35)$$

This, however, is precisely the formula for standard adjustment, corresponding to normal equations formed after the initial iteration. In particular, the left-hand side is $\Delta(\Delta u_\beta)$, and the equation corresponds to the system (8) considered at the point (P). We conclude that in this theoretical case, which includes the category of error-free observations, the standard solution is valid to the third order and hence conforms to the strategy pursued in this study.

The second case, which satisfies the standard assumption that both the parametric corrections and the residuals are small (first-order), is certainly the most significant in practice. As was indicated at the close of Section 3.1, equation (23a) becomes

$$[(a_{\beta\alpha}) - (\delta x_s) \Omega_{\beta\alpha}^s] \Delta(\Delta u^\alpha) = (A_\beta^s) (\delta x_s) . \quad (36)$$

In comparing it with the standard equation of the form (35), we notice that the latter entails the error $(\delta x_s) \Omega_{\beta\alpha}^s \Delta(\Delta u^\alpha)$, which is of the third order. The standard formula in such practical cases is thus valid only to the second order. The newly developed equation (36) is superior to the standard normal equations by one order of magnitude, which leads to a commensurate increase in the rate of convergence. In fact, depending on the smallness of Δu^α , and thereby on the proximity of the initial point P to the least-squares point \bar{P} , the solution of (36) may be final.

Next, we examine the improvement in the solution if equation (36) is re-applied. We further assume that the original solution $\Delta(\Delta u^\alpha)$ is valid to the third order, the same as the accuracy of the formula (36) itself. The new, updated set (Δu^α) is thus valid to the third order (instead of the first order as assumed initially), and the new $\Delta(\Delta u^\alpha)$ contains fourth-order quantities. At this stage (36) reveals that fifth-order quantities are retained, while sixth- and higher-order quantities are ignored. The new solution will thus be valid to the fifth order. By contrast, the second iteration in the standard approach (i.e., the original application of equation 35) would yield the solution valid to the second order, the next iteration would yield the solution valid to the third order, and so on. To summarize, the order of errors in the repeated applications of (36) would be 4, 6, 8, ... , while this order in the standard iterative process symbolized by (35) would be 3, 4, 5,

From the theoretical standpoint, the most revealing is the third case. As we have seen, the original formulation (23a) is valid to within the third order. However, if δx^r contains finite quantities, the standard form of normal equations (35) is virtually useless; it is valid only to the first order, since the first neglected term in (23a), namely $-(\delta x_s) \Omega_{\beta\alpha}^s \Delta(\Delta u^\alpha)$, is of the second order, the same as the entire left-hand side of (35). Thus, not only does (35) entail errors two orders of magnitude greater than errors in (23a), but, in general, it does not converge. (If δx^r should belong to an "in-between"

category discussed in Section 3.2, these statements would be commensurably softened.) Such facts are at the root of the standard assumption that not only Δu^α and thereby (Δu^α) , but also δx^r and thereby (δx^r) should be small. We conclude by stating that the standard approach cannot accommodate the third case, while the new formula (23a) accommodates it to the third order, and the simplified formula (36) accommodates it to the second order.

As a matter of interest, we examine the possibility of $\Delta(\Delta u^\alpha)$ being larger than second-order. Out of necessity, we assume that $\Delta(\Delta u^\alpha)$ contains quantities whose magnitude is somewhere between the first and the second orders, in the sense that $o[\Delta(\Delta u^\alpha) \Delta(\Delta u^\beta)]$ is of the third order (instead of the fourth order as stipulated earlier). Clearly, the neglected quantities $o[\Delta(\Delta u^\alpha) \Delta(\Delta u^\beta)]$ cannot be of the second order since this would make even the full equation (23a) valid only to the first order. The dots in equations (17b) and (19b) now represent third- (instead of fourth-) and higher-order quantities. The same can be said about the terms a, b, and c, and thereby about (22) and about errors in (23a). Due to the modification of the assumption regarding $\Delta(\Delta u^\alpha)$, the accuracy of the three formulas under consideration (equations 35, 36, and 23a) is somewhat compromised in the first and the second cases, but the effect is not qualitatively profound. However, in the third case the standard formula (35) is virtually useless. The simplified formula (36), although compromised to a certain extent, leads to useful results, while the full formula (23a) is as effective as the standard formula (35) would be under the standard assumptions. (Here again, if δx^r should belong to an "in-between" category, all three formulas would give commensurably better results.)

We note that (23a) could be reformulated with the partial derivatives evaluated at (P). Since at (P) we have

$$(\Omega_{\beta\alpha}^s) = \Omega_{\beta\alpha}^s + \Phi_{\beta\alpha\gamma}^s (\Delta u^\gamma) + \dots, \quad (37a)$$

in keeping with the stipulated error level we can write (23a) as

$$[(a_{\beta\alpha}) - (\delta x_s) (\Omega_{\beta\alpha}^s)] \Delta(\Delta u^\alpha) = (A_\beta^s) (\delta x_s). \quad (37b)$$

However, the original equation (23a) is preferable to (37b) since it displays the fact that the sets $\Omega_{\beta\alpha}^s$ and $\Phi_{\beta\alpha\gamma}^s$ can be stored and treated as constant, and implicitly suggests in which manner and under what conditions it can be simplified. We have seen such simplifications in (35) and (36). Although the

full formula (23a) is appealing from the theoretical standpoint, computer run-time and storage considerations favor the simplified formula (36).

After obtaining the final parametric values $u^\alpha = u_0^\alpha + \Delta u^\alpha$, one can compute the desired quantities, such as Δx^r and thereby $\delta \bar{x}^r$, as well as the tensor \bar{A}_α^r needed to form

$$\bar{a}_{\beta\alpha} = \bar{A}_\beta^s g_{sr} \bar{A}_\alpha^r, \quad (38)$$

which, in turn, can be used in the variance-covariance propagation and in the weight propagation. In paralleling the discussion of Section 3.1 pertaining to a linear model, we state that the tensor $\bar{a}^{\alpha\beta}$, obtained from $\bar{a}_{\beta\alpha}$ via the tensor equation $\bar{a}^{\alpha\beta} \bar{a}_{\beta\gamma} = \delta_\gamma^\alpha$, corresponds to the variance-covariance matrix of adjusted parameters, while the tensor $\bar{a}_{\beta\alpha}$ itself corresponds to the weight matrix of these parameters. Further, the variance-covariance and the weight matrices of adjusted observations correspond respectively to

$$\bar{g}^{rs} = \bar{A}_\alpha^r \bar{a}^{\alpha\beta} \bar{A}_\beta^s, \quad \bar{g}'_{sr} = g_{si} \bar{g}^{ij} g_{jr}. \quad (39a,b)$$

As has been mentioned at the close of Chapter 2, the above tensors describe "tangential", or first-order, properties of least squares estimates. In reference to [Blaha, 1984], the variance-covariance and the weight matrices of residuals correspond respectively to the tensors

$$\bar{g}''^{rs} = g^{rs} - \bar{g}^{rs}, \quad \bar{g}''_{sr} = g_{sr} - \bar{g}'_{sr}. \quad (40a,b)$$

4.2 Additional Notes

We begin this discussion by demonstrating that an eventual re-application of (23a) is straightforward and without any pitfalls. We observe that such a re-application would be equivalent to using (23a) in conjunction with a new point (P) situated closer to \bar{P} than the original point (P) obtained via the initial iteration. The quantity (Δu^r) would then represent the model-surface coordinate differences between the new point (P) and the initial point P; (δx^r) would similarly refer to the new point (P). This situation would change nothing in the development nor in the results represented by equation (23a) itself. In particular, the utilization of (Δu_β) from (8), which in Fig. 1 gave rise to the model-plane point P_1 and thereby to the model-surface point (P), was essentially

a matter of convenience. If such a model-plane point were found closer to P_2 (produced e.g. by the initial iteration of Section 3.2), it would represent a welcome shortcut, but none of the pertinent relationships from (12a) on would change. And if it could be found at the exact location of P_2 , the right-hand side of (23a) would be zero by virtue of (7), which would yield $\Delta(\Delta u^\alpha) = 0$. The same considerations are valid also with regard to the simplified formula (36). As we have seen, the latter differs from the full formula (23a) by the error level due to the truncated terms, but only in the case of large residuals.

We now focus our attention on the second case (with small residuals) represented by (36), and examine whether advantage can be taken of the set $\Omega_{\alpha\beta}^r$ in forming (A_α^r) and (δx^r) , as opposed to computing these tensors from the nonlinear model. If we include only the term containing $\Omega_{\alpha\beta}^r$ in (16) and neglect the subsequent terms, the error in (A_α^r) will be of the second order, as will be the error in $(a_{\beta\alpha})$ formed through (23b). Accordingly, the error on the left-hand side of (36) will be of the fourth order and negligible. Similarly, if we include only the term containing $\Omega_{\alpha\beta}^r$ in (18) and neglect the subsequent terms, the error in (Δx^r) and thereby also in (δx^r) given by (20b) will be of the third order, resulting only in a fifth-order error on the left-hand side of (36).

However, the right-hand side of (36) offers a different picture. In forming $(A_\beta^s)(\delta x_s)$, where (A_β^s) contains second-order errors and the first-order quantity (δx_s) contains third-order errors, we observe that either kind of errors introduces third-order errors into the product. Thus, to keep with the fourth-order error level implied by (36), both (A_α^r) from (16) and (Δx^r) from (18) would have to include the terms with $\Phi_{\alpha\beta\gamma}^r$. We note that the order of magnitude of the remaining truncation errors in these two quantities is independent of the closeness of (Δu^α) to Δu^α , which would preclude a re-application of (36).

In the case of a single application of (36), only (16) would entail significant computing effort since (18) could take advantage of the intermediate results $\Omega_{\alpha\beta}^r(\Delta u^\beta)$ and $\Phi_{\alpha\beta\gamma}^r(\Delta u^\beta)(\Delta u^\gamma)$. If one chose to pursue this avenue, it would then be quite easy to adopt the full formula (23a) where $\Phi_{\beta\alpha\gamma}^s(\Delta u^\gamma)$ would be known from an early intermediate result. However, passing from (36) to (23a) would not decrease the overall error level (fourth-order). In realizing that $\Phi_{\alpha\beta\gamma}^r$ is not needed in (36), we might wish to avoid computing and storing these

third-order partial derivatives altogether. Accordingly, (A_α^r) as well as (Δx^r) and thereby $(\delta x^r) = \delta x^r - (\Delta x^r) \equiv x_Q^r - (x^r)$, where $(x^r) = x^r((u^\alpha))$, would be computed from the nonlinear model as in the standard adjustment approach. This choice is further supported by a potential need to re-apply the algorithm (36).

It is instructive to compare the development presented in this study with that of [Blaha, 1987]. Although both cases rely on the precise relation (7), the latter resulted in a less efficient algorithm due to the simplification which we shall now describe. The tensor \bar{A}_β^s was utilized in the form $A_\beta^s + \Delta A_\beta^s$ as in (13a). On the other hand, the vector $\delta \bar{x}^r$ was substituted for by the right-hand side of (14), but only for the contraction with $A_\beta^s g_{sr}$, not for the contraction with $\Delta A_\beta^s g_{sr}$. Accordingly, the relation (7) became

$$A_\beta^s g_{sr} \delta x^r - A_\beta^s g_{sr} \Delta x^r + \Delta A_\beta^s g_{sr} \delta \bar{x}^r = 0.$$

The first term above was (Δu_β) by virtue of (8), and the second term was $-a_{\beta\gamma} \Delta u^\gamma - A_\beta^s g_{sr} v^r$, as is confirmed through (12a,b). The resulting equation was contracted with $a^{\alpha\beta}$, yielding

$$\Delta u^\alpha = (\Delta u^\alpha) - a^{\alpha\beta} A_\beta^s g_{sr} v^r + a^{\alpha\beta} \Delta A_\beta^s g_{sr} \delta \bar{x}^r,$$

which, except for minor notational differences, was presented in [Blaha, 1987] as equation (28).

The solution in this reference proceeded as if Δu^α were split into $(\Delta u^\alpha) + \Delta(\Delta u^\alpha)$ seen in (15), but only for the left-hand side above, not for the formation of v^r by (12c), nor for the formation of $\delta \bar{x}^r$ by (14) with Δx^r given by (12a-c). The values Δu^α implied in v^r and Δx^r were approximated by (Δu^α) , i.e., the terms beyond the first on the right-hand side of the above equation were treated as if $\Delta(\Delta u^\alpha)$ were zero. This introduced errors in an initially exact equation, which were an order of magnitude greater than the terms neglected in the present study. For example, in the situation called here the second case the thus introduced errors would be of the third order instead of the current fourth order. Such simplifications downgraded the convergence characteristics of [Blaha, 1987] essentially to the level of the standard approach.

The results as well as intermediate formulas in this study could be presented in the familiar matrix notation. In the case of simply- and doubly-indexed quantities, matrix transcriptions would present no difficulty since

tensor contractions and matrix multiplications amount to the same operations, provided the matrices are arranged in the proper sequence and form (direct or transposed). With regard to three- and higher-dimensional arrays, matrix manipulations could be extended upon imagining that ordinary matrices are stacked one behind the other, that groups of matrices are stacked one behind the other, etc. However, we prefer working with the original indexed quantities, which allows more flexibility in derivations as well as computations, and which, in any case, is at the root of the matrix operations themselves (restricted to one or two indices). For the sake of interest, Appendix D translates the main outcome of this study into matrix notation.

5. SUMMARY AND CONCLUSIONS

A convenient approach for resolving nonlinear least-squares problems consists in using an isomorphic geometrical setup with tensor structure and notation. Such a link is highlighted by the consideration that if, in the n -dimensional observational space, the metric tensor g_{sr} corresponds to the weight matrix P , and if the contravariant components $\delta\bar{x}^r$ of a displacement symbolized by $\delta\bar{x}$ correspond to a set of estimated observational errors, i.e., to a set of negative residuals $-V$, then the least-squares criterion

$$V^T P V = \text{minimum}$$

corresponds to the quadratic form

$$\delta\bar{x}^s g_{sr} \delta\bar{x}^r \equiv \delta\bar{s}^2 = \text{minimum},$$

where $\delta\bar{s}$ is the magnitude of the displacement. Should $\delta\bar{x}$ identify an unknown displacement from the u -dimensional model surface to the point Q given in the surrounding observational space through the actual observations, the minimum-distance property restricts $\delta\bar{x}$ to the segment $\bar{P}Q$ of a straight line dropped orthogonally from Q onto the model surface, where \bar{P} marks the intersection. This situation is illustrated schematically in Fig. 1. Since orthogonal relations are readily exploited in a geometrical environment, the least-squares method -- unlike any other method -- is particularly suited for geometrical analogy and analysis.

Among the basic correspondences between adjustments and geometry, the number of observations, n , and the number of parameters, u , correspond respectively to the dimensionality of the observational space and of the model surface. The variance-covariance matrix of observations, Σ , corresponds to the associated metric tensor g^{rs} , while the weight matrix of observations, adopted as $P = \Sigma^{-1}$, corresponds to the metric tensor g_{sr} . Since Σ and thus also P are constant, the observational space is endowed with a coordinate system $\{x^r\}$ such that

$$g_{sr} = \text{constant}, \quad g^{rs} = \text{constant}.$$

The set L^b of actual observations corresponds to the set x_Q^r of observational-space coordinates describing the point Q . All possible sets L^a of adjusted

observations (subject to no criterion) correspond to the Gauss form of the model surface endowed with a coordinate system $\{u^\alpha\}$:

$$x^r = x^r(u^\alpha) , \quad r = 1, 2, \dots, n , \quad \alpha = 1, 2, \dots, u .$$

The final set of adjusted parameters, X^a , corresponds to a particular set u^α of model-surface coordinates describing the least-squares point \bar{P} . The set of initial parameters, X^0 , corresponds to the set u_0^α of model-surface coordinates describing the initial point P . The final set of parametric corrections, X , then corresponds to $\Delta u^\alpha = u^\alpha - u_0^\alpha$; these quantities are assumed to be small (termed first-order). The final set of adjusted observations, L^a , corresponds to a particular set $x^r = x^r(u^\alpha)$ of observational-space coordinates describing the least-squares point \bar{P} . Similarly, the set $P(X^0)$ corresponds to the set $x_0^r = x^r(u_0^\alpha)$ of observational-space coordinates describing the initial point P . The set of negative constant terms, $-L = L^b - F(X^0)$, corresponds to the contravariant vector $\delta x^r = x_Q^r - x_0^r$, while the set of negative residuals, $-V = L^b - F(X^a)$, corresponds to the contravariant vector $\delta \bar{x}^r = x_Q^r - x^r$. The initial design matrix A , which in standard observation equations $V = AX + L$ relates the parametric corrections to the residuals, corresponds to the design tensor $A_\alpha^r = \partial x^r / \partial u^\alpha$ evaluated at P . On the other hand, the standard adjustment approach does not have equivalents of $\Omega_{\alpha\beta}^r$ and $\Phi_{\alpha\beta\gamma}^r$, which form three- and four-dimensional arrays, respectively, and contain the second- and the third-order partial derivatives of x^r with respect to u^α , evaluated at P .

The geometrical approach is based on a direct exploitation of the relation

$$\bar{A}_\beta^s g_{sr} \delta \bar{x}^r = 0 , \quad (41)$$

where \bar{A}_β^s represents the design tensor evaluated at the least-squares point \bar{P} , and equation (41) itself represents the orthogonality condition at \bar{P} . The outcome of the geometrical development is considered in two methods, called geometrical and extended geometrical. It is contrasted to the standard method treating nonlinear models in a linearized form. The algorithms associated with all three methods are presented below in the form of the first iteration, and in the form of the second and subsequent iterations.

All the formulas are written in tensor notation. In the case of simply- and doubly-indexed quantities, matrix transcriptions would present no difficulty since tensor contractions and matrix multiplications amount to the same

operations, provided the matrices are arranged in the proper sequence and form (direct or transposed). With regard to three- and higher-dimensional arrays, matrix manipulations could be extended upon imagining that ordinary matrices are stacked one behind the other, that groups of matrices are stacked one behind the other, etc. However, working with the original indexed quantities allows more flexibility in derivations as well as in computations.

In tensor notation, the initial matrix of normal equations corresponds to the model-surface metric tensor $a_{\beta\alpha}$ at the initial point P, and the initial right-hand side of normal equations corresponds to the model-surface covariant vector $A_{\beta}^s \delta x_s$ at P, where

$$a_{\beta\alpha} = A_{\beta}^s g_{sr} A_{\alpha}^r, \quad \delta x_s = g_{sr} \delta x^r.$$

The parametric corrections stemming from the first iteration are symbolized by (Δu^{α}) , and they give rise to an updated point (P). The latter is described by the model-surface coordinates $(u^{\alpha}) = u_0^{\alpha} + (\Delta u^{\alpha})$. The quantities belonging to (P) will likewise be written in parentheses. The parametric corrections obtained in the second iteration will be denoted $\Delta(\Delta u^{\alpha})$, and they will give rise to a new updated point determined via $(u^{\alpha}) + \Delta(\Delta u^{\alpha})$. The notation used in conjunction with the second iteration will be retained also for any further iterations.

Standard method. Under the assumption that both sets Δu^{α} and δx^r contain small quantities (first-order), the first iteration in the standard method reads

$$a_{\beta\alpha} (\Delta u^{\alpha}) = A_{\beta}^s \delta x_s, \quad (42a)$$

representing the initial normal equations. The second and further iterations follow the same principle:

$$(a_{\beta\alpha}) \Delta(\Delta u^{\alpha}) = (A_{\beta}^s) (\delta x_s). \quad (42b)$$

As is illustrated by the example in Appendix A, the convergence properties of this method may be significantly impaired by an increased size of observational errors.

Geometrical method. Under the same assumption as above (both sets Δu^{α} and δx^r contain small quantities), the first iteration utilizes the same formula as its standard counterpart, namely

$$a_{\beta\alpha} (\Delta u^{\alpha}) = A_{\beta}^s \delta x_s. \quad (43a)$$

However, the second and further iterations proceed according to

$$[(a_{\beta\alpha}) - (\delta x_s) \Omega_{\beta\alpha}^s] \Delta(\Delta u^\alpha) = (A_\beta^s) (\delta x_s), \quad (43b)$$

representing the modified normal equations. The triply-indexed quantity $\Omega_{\beta\alpha}^s$, formed by second-order partial derivatives of the observables with respect to the parameters, is evaluated only at the initial point P. The formulas (43a,b) are relatively simple yet efficient, and are suitable for practical applications where the convergence of the standard method is problematic.

Extended geometrical method. Although the assumption regarding Δu^α is unchanged, this method is tailored for δx^r containing relatively large quantities, for which the first iteration reads

$$[a_{\beta\alpha} - \delta x_s \Omega_{\beta\alpha}^s] (\Delta u^\alpha) = A_\beta^s \delta x_s. \quad (44a)$$

Compared to the geometrical method, the current algorithm is seen to utilize second-order partial derivatives and to give rise to the modified normal equations already in its first iteration. The formula for the second and further iterations is given as

$$[(a_{\beta\alpha}) - (\delta x_s) \Omega_{\beta\alpha}^s - (\delta x_s) \Phi_{\beta\alpha\gamma}^s (\Delta u^\gamma)] \Delta(\Delta u^\alpha) = (A_\beta^s) (\delta x_s), \quad (44b)$$

representing the modified normal equations at updated stages. Here use is made of $\Phi_{\beta\alpha\gamma}^s$ containing third-order partial derivatives. This quantity is evaluated only at P, similar in this respect to $\Omega_{\beta\alpha}^s$. We note that the quantity inside the brackets of (44b) could be replaced by $(a_{\beta\alpha}) - (\delta x_s) (\Omega_{\beta\alpha}^s)$, where $(\Omega_{\beta\alpha}^s)$ would represent the second-order partial derivatives evaluated at an updated point.

After obtaining the final set of parametric values $u^\alpha = u_0^\alpha + \Delta u^\alpha$, one can compute the desired quantities \bar{A}_α^r , Δx^r , and δx^r either from the nonlinear model or by using the Taylor series expansion. The tensor \bar{A}_α^r is needed to form

$$\bar{a}_{\beta\alpha} = \bar{A}_\beta^s g_{sr} \bar{A}_\alpha^r,$$

which corresponds to the weight matrix of adjusted parameters. The tensor $\bar{a}^{\alpha\beta}$ obtained from $\bar{a}_{\beta\alpha}$ through the relation $\bar{a}^{\alpha\beta} \bar{a}_{\beta\gamma} = \delta_\gamma^\alpha$ then corresponds to the variance-covariance matrix of adjusted parameters. Further we have

$$\bar{g}^{rs} = \bar{A}_\alpha^r \bar{a}^{\alpha\beta} \bar{A}_\beta^s, \quad \bar{g}_{sr} = g_{si} \bar{g}^{ij} g_{jr},$$

which correspond respectively to the variance-covariance and the weight matrices of adjusted observations. The above tensors are expressible by means of orthonormal vectors tangent to the model surface at \bar{P} . Accordingly, they describe "tangential", or first order, properties of least-squares estimates. Finally, the tensors

$$\bar{g}''^{rs} = g^{rs} - \bar{g}'^{rs}, \quad \bar{g}''_{sr} = g_{sr} - \bar{g}'_{sr}$$

correspond respectively to the variance-covariance and the weight matrices of residuals.

A significant contribution to the understanding of the least-squares adjustment theory through differential geometry has been made by Teunissen [1985]. The geometry of a nonlinear adjustment is treated in his Chapter IV, the main topic of which is Gauss' iteration method. This method characterizes the standard adjustment approach as is confirmed, for example, by the text on page 109 stating that "at each iteration step the observation point is orthogonally projected onto a new tangent space". Such an algorithm is equivalent to the solution of the type (42a,b) above, which, in terms of Fig. 1, would result in the projection of the point Q onto the model plane passing through the initial point P, followed by the projection of Q onto a new model plane passing through an updated point (P), etc.

In the same vein, page 109 (ibid.) states that Gauss' iteration method "consists of successively solving a linear least-distance adjustment until the necessary condition of orthogonality is fulfilled". This condition, equivalent to our equation (41), is fulfilled in the standard method as a by-product of the projections discussed above. By contrast, the geometrical approach actively seeks to fulfill it at every step. A one-step solution producing the least-squares point \bar{P} directly is hindered only by the necessity to truncate certain terms, but not to the extent of making the entire model linear (see the above equations 43b and 44a,b). The matrix of modified normal equations generated in the process is positive-definite, similar in this respect to the matrix of normal equations in the standard method.

Although the behavior of the standard adjustment algorithm has been analyzed and described in detail by Teunissen [1985], the behavior of the geometrical algorithm developed herein allows ample opportunity for further analysis. A study comparing the geometrical approach with alternative methods

would be worthwhile for the sake of theory as well as for practical reasons (computational burden, rate of convergence, etc.). Another worthwhile effort would be to study the effect of the model's nonlinearity on the statistical properties of least-squares estimators, as is outlined in Section 6 of Chapter IV, *ibid.* The need for further research is also underscored by the fact that the above discussion has touched merely on one class of adjustment problems, the parametric adjustment. Nonlinear versions of the general method (with or without constraints), rank-deficient systems, etc., all await a systematic geometrical treatment.

Indeed, as the Epilogue to his Chapter IV, Teunissen [1985] indicates that scientists are only on the brink of understanding the complexity of nonlinear adjustment. He further states: "Unfortunately, one can seldom extend the elegant formulation and solution techniques from linear to non-linear situations. For most non-linear problems one will therefore have to recourse, in practice, to methods which are iterative in nature". In this light, the present study constitutes a contribution in accelerating, although not completely eliminating, the iterative process.

Encouraging results have been obtained in the numerical example presented in Appendix A, illustrating convergence properties of a third-order polynomial in four variables. Although the standard method converged slowly in one of four analyzed cases and diverged in two others, the geometrical method converged in two and three iterations, respectively. The extended geometrical method further reduced the number of iterations from three to two. It is expected that in most nonlinear cases the presence of second-order partial derivatives will translate into two iterations in the geometrical method as compared to several iterations needed by the standard method.

In spite of the accelerated convergence of the geometrical approach, the fact remains that the methods discussed herein, as well as other methods used in physical sciences for adjusting nonlinear models, are iterative in nature. However, one can conceive of a geometrical scheme where the determination of the least-squares point \bar{P} can be made arbitrarily accurate in one single step upon increasing the number of meaningful terms in the Taylor series expansions. Such an approach would have to be fundamentally different from the one described in this study, where an increase in the number of meaningful terms has certain limitations. In particular, if the observational errors are relatively small,

the current geometrical approach cannot take advantage of the third- and higher-order partial derivatives of the observables with respect to the parameters, whose contribution would be invalidated by the truncation errors of comparable magnitudes. Thus, developing a theoretically distinct geometrical setup along these lines is yet another challenge in the field of the nonlinear least-squares adjustment.

APPENDIX A

NUMERICAL EXAMPLE

The nonlinear parametric adjustment model is represented by six observables, each expressed by a third-order polynomial in four variables. Thus, the number of observations is six, $n=6$, and the number of parameters is four, $u=4$. A simpler adjustment problem featuring a second-order polynomial was presented by Bessette [1987], who utilized the algorithm described in [Blaha, 1987]. Although that algorithm lowered the number of scalar multiplications in each iteration as compared to the standard method, it did not accelerate the rate of convergence (see Section 4.2). It has been superseded by the algorithm developed here, which has been programmed and verified independently on two different computers.

In terms of the four parameters u^α , $\alpha=1,2,3,4$, a general observable denoted x is expressed by

$$x = c_\alpha u^\alpha + k_{\alpha\beta} u^\alpha u^\beta + m_{\alpha\beta\gamma} u^\alpha u^\beta u^\gamma,$$

where α ranges from 1 to 4, β ranges from α to 4, and γ ranges from β to 4. The numerical values of the four c -coefficients (linear), the ten k -coefficients (quadratic), and the twenty m -coefficients (cubic) for all six observables are:

+0.8, +1.2, +0.8, +1.5; +2.3, -1.6, +1.7, +2.0, +1.8, -2.1, +2.1, +1.5,
+2.4, +2.1; -1.1, +0.8, -0.9, -2.1, +1.0, -2.0, +1.5, +2.0, -0.9, -1.2, +0.8,
-1.9, -1.3, +0.9, +1.2, +0.8, -2.1, -1.6, +1.8, +1.1.

+1.7, +2.1, -2.1, +1.6; +1.3, +2.2, +1.0, -2.2, +1.7, +1.7, +2.1, +1.8,
+0.7, +1.6; +1.3, -0.8, -0.9, -2.1, +2.0, +1.2, -1.8, +0.9, -1.3, +1.3, -2.1,
-0.9, +1.5, +0.8, +1.1, +1.7, -2.1, -1.8, +1.2, -1.0.

+1.3, -2.1, +1.7, +2.1; +2.2, +2.3, -2.2, +1.1, +1.2, +2.0, -1.7, +1.5,
+1.9, +1.7; -1.1, -2.0, +1.7, +0.8, +1.0, -2.1, +2.2, +1.2, -1.9, +1.7, -1.4,
-0.8, +0.8, +1.0, -1.1, -0.9, +1.8, -1.4, -0.9, +2.1.

+1.6, +1.0, -2.1, +1.1; -1.7, +2.2, +1.6, -1.6, +1.6, -1.2, +2.1, +2.3,
-2.1, +1.2; -2.2, -0.8, +1.0, -1.2, +0.9, +1.3, -0.8, +2.1, +1.9, +1.2, +0.9,
-1.5, -1.6, -0.8, -1.0, +0.9, -0.9, +2.1, +0.8, -1.3.

+0.9, +2.1, -2.1, +1.8; +1.2, +1.7, +2.2, -1.5, +1.9, +1.3, -2.2, +1.7,
+1.0, +2.3; +1.1, -1.2, -0.9, +1.3, +1.5, +1.0, -0.8, -2.2, +0.8, +1.7, -1.2,
-0.3, +1.1, -2.0, +1.7, -0.9, +0.8, +1.3, -1.8, -1.1.

+1.7, +1.2, +1.7, -1.6; +2.1, +1.6, -1.1, +1.9, +1.6, +1.7, +2.0, +1.2,
+1.6, -1.7; +1.3, +2.1, -1.0, -0.8, +1.2, -2.1, +2.2, -1.6, -1.8, -0.8, +1.9,
+1.2, -0.9, +1.2, +2.0, -1.7, +1.1, +1.0, +0.8, -2.2.

The positive-definite variance-covariance matrix of observations corresponds to the associated metric tensor g^{rs} , chosen as

$$g^{rs} = \begin{array}{cccccc} +1.3266 & +0.5254 & +0.3669 & +0.1572 & -0.2667 & -0.0549 \\ . & +1.3457 & -0.0281 & +0.5924 & -0.1448 & +0.5521 \\ . & . & +1.4143 & +0.5144 & +0.3600 & -0.5204 \\ . & . & . & +1.4629 & -0.5604 & +0.4877 \\ . & . & . & . & +1.3930 & -0.1034 \\ . & . & . & . & . & +1.3500 \end{array}$$

where the dots denote the symmetric elements. The positive-definite weight matrix of observations corresponds to the metric tensor g_{sr} computed from $g^{ks} g_{sr} = \delta_r^k$. The initial set of parameters is represented by

$$u^\alpha = +1.399, -1.201, +1.299, +1.099,$$

describing the initial point P by model-surface coordinates. The observational-space coordinates x_0^r of this point are determined from the model, as are the sets of partial derivatives A_α^r , $\Omega_{\alpha\beta}^r$, and $\Phi_{\alpha\beta\gamma}^r$ (the last set is constant here, and any further set is zero). Finally, the elements of x^r representing error-free observations are

$$x^r = +4.80, -4.79, +33.44, +0.33, +22.47, -11.69.$$

These values correspond to the following set of final parameters:

$$u^\alpha = +1.4, -1.2, +1.3, +1.1.$$

The example has been generated in four versions which differ from one another by the size of the final residuals. Each version is treated by three methods summarized in the conclusion: standard, geometrical, and extended geometrical. A basic version has been generated with excessively large final

residuals grouped in the tensor $-\delta\bar{x}^F$:

$$\delta\bar{x}^F = +3.22, +22.55, -10.59, -1.28, +24.27, +14.68 .$$

This version gives rise to a total of four, upon scaling the values in $\delta\bar{x}^F$ by a factor c , where

$$c = 0.001; 0.01; 0.1; 1.$$

The residuals in the first version ($c=0.001$) can be considered first-order, or essentially differential, as required by the standard methodology. On the other side of the spectrum, the residuals in the fourth version ($c=1$) are so large that this case is presented mainly for the sake of theoretical interest.

The iterative process is terminated in each method and each version when errors in all of the parameters u^α , $\alpha=1,2,3,4$, are less or equal to 5×10^{-8} . Such errors remaining in u^α are listed below only for u^2 , for which they are in general the worst. In the standard method, the errors in u^4 are very similar to those in u^2 (including the sign), but their magnitude is slightly larger. A statement of this kind applies also to the geometrical method for small values of c (represented by $c=0.001$ and $c=0.01$). The reverse is generally true for large values of c (represented by $c=0.1$ and $c=1$) in the geometrical method, and for all four c 's in the extended geometrical method.

Results for u^2 in the standard method, in the geometrical method, and in the extended geometrical method are:

$$c = 0.001 \quad \text{..... stand.: } -4 \times 10^{-6}, +3 \times 10^{-8} ;$$

$$\text{geom.: } -4 \times 10^{-6}, +2 \times 10^{-12} .$$

$$c = 0.01 \quad \text{..... stand.: } -3 \times 10^{-5}, +2 \times 10^{-6}, -3 \times 10^{-7}, +3 \times 10^{-8} ;$$

$$\text{geom.: } -3 \times 10^{-5}, +2 \times 10^{-9} .$$

$$c = 0.1 \quad \text{..... stand.: } -2 \times 10^{-4}, +2 \times 10^{-4}, -3 \times 10^{-4}, +3 \times 10^{-4}, \text{ diverges ;}$$

$$\text{geom.: } -2 \times 10^{-4}, +1 \times 10^{-7}, -1 \times 10^{-10} ;$$

$$\text{ext. geom.: } -3 \times 10^{-5}, -1 \times 10^{-10} .$$

$c = 1$ stand.: -2×10^{-3} , $+2 \times 10^{-2}$, -2×10^{-1} , $+1 \times 10^0$, diverges ;
 geom.: -2×10^{-3} , -1×10^{-5} , $+1 \times 10^{-8}$;
 ext. geom.: -1×10^{-5} , -2×10^{-10} .

The above results and further computer runs warrant the following comments:

1) Results substantially worse than those listed for u^2 have been noticed for $c=0.001$ in the case of the parameter u^4 in the geometrical method ($+1 \times 10^{-11}$ as opposed to $+2 \times 10^{-12}$). However, such exceptions do not detract from the overall pattern seen above.

2) The passage from the geometrical method to the extended geometrical method, i.e., the inclusion of second-order partial derivatives in the first iteration and of third-order partial derivatives in the second iteration, has no appreciable effect on the adjustment with small residuals ($c=0.001$ and $c=0.01$). On the other hand, in case of large residuals ($c=0.1$ and $c=1$) the effect is seen to be great and the number of iterations is reduced from three to two.

3) If, in the extended geometrical method for $c=0.1$ and $c=1$, the first iteration is adopted from the standard method (i.e., if the second-order partial derivatives are excluded from that iteration), the total number of iterations will grow from two to three. As an example in the case $c=1$, the errors in the third iteration in such a mixed method have the same order of magnitude as the errors in the second iteration of the extended geometrical method proper.

4) If, in the extended geometrical method for $c=0.1$ and $c=1$, the second iteration uses only second-order partial derivatives (in keeping with the pattern of the first iteration), the errors will worsen by approximately two orders of magnitude. However, this does not lead to an increase in the number of iterations (namely two) resulting from the current cut-off criterion.

5) In this example, where fourth- and higher-order partial derivatives are zero, a procedure utilizing updated second-order partial derivatives in the second iteration of the geometrical method agrees perfectly with the extended geometrical method. This useful verification can be described as follows. In general, the updated second-order partial derivatives can be expressed by

$$(\alpha_{\beta\alpha}^s) = \alpha_{\beta\alpha}^s + \phi_{\beta\alpha\gamma}^s (\Delta u^\gamma) + \dots$$

Since here the quantities represented by the dots are zero, the equality holds true with only two terms on the right-hand side. But this is precisely the form featured by the extended geometrical method.

6) In the geometrical and the extended geometrical methods, as well as in the mixed methods of the comments numbered 3 and 4, the solution of the modified normal equations has been carried out using both the Gauss elimination algorithm and the Choleski algorithm for positive-definite matrices. All the results, including the inverted matrices of modified normal equations in all iterations, have been found identical to all 11 digits printed. This illustrates that in the usual case of continuous second-order partial derivatives and of the solution converging to a minimum, the matrix of modified normal equations evaluated in the neighborhood of the final point \bar{P} is positive-definite.

The above example illustrates that the standard method converges in general only for small residuals (here for $c=0.001$ and $c=0.01$). When larger residuals are present, this method either converges very slowly or diverges (the latter has been observed for $c=0.1$ and $c=1$). The geometrical method is particularly attractive. It is relatively simple, yet it converges in two iterations for small residuals ($c=0.001$ and $c=0.01$), and in three iterations for large residuals ($c=0.1$ and $c=1$). In the case of large residuals, the use of the extended geometrical method further reduces the number of iterations to two. This reduction has been accomplished here even without third-order partial derivatives in the second iteration, which suggests that a fourth method could be adopted (the mixed method of the comment 4), where only second-order partial derivatives evaluated at P would be used, whether in the first or in subsequent iterations. In conclusion, we have seen that the standard procedure may be slow to converge, or may diverge if the residuals are large, whereas the use of geometry leads to fast convergence, represented here by two or three iterations depending on the level of simplifications.

APPENDIX B

MODEL SURFACE AND MODEL PLANE

In this study, the underlying space encompassing all the geometrical objects is an n -dimensional flat space called the "observational space". The space coordinate system used exclusively is symbolized by $\{x^r\}$, $r=1,2,\dots,n$, in which the metric tensor is

$$g_{sr} = \text{constant} . \quad (B.1)$$

In matrix notation, g_{sr} would be represented by a positive-definite matrix of dimensions $(n \times n)$. Embedded in the observational space is a u -dimensional surface called the "model surface". The surface coordinate system is symbolized by $\{u^\alpha\}$, $\alpha=1,2,\dots,u$. The surface coordinates are considered in the role of u independent variables, without any specific physical meaning.

In the observational space, the model surface is defined by the Gauss form:

$$x^r = x^r(u^\alpha) , \quad r = 1,2,\dots,n , \quad \alpha = 1,2,\dots,u . \quad (B.2)$$

Equation (B.2) restricts the values the coordinates x^r can have in order to describe points on the model surface. This form, expressing each of the n space coordinates as some function of the u surface coordinates, is considered known. It is a higher-dimensional analogue of the Gauss form presented in Chapter 6 of [Hotine, 1969], describing a two-dimensional surface ($u=2$) embedded in a three-dimensional space ($n=3$).

In the vicinity of an "initial" point P , whose model-surface coordinates are u_0^α , (B.2) can be expanded in the Taylor series:

$$\begin{aligned} x^r = x^r(u^\alpha) = & x_0^r + A_\alpha^r (u^\alpha - u_0^\alpha) + (1/2) \Omega_{\alpha\beta}^r (u^\alpha - u_0^\alpha)(u^\beta - u_0^\beta) \\ & + (1/6) \Phi_{\alpha\beta\gamma}^r (u^\alpha - u_0^\alpha)(u^\beta - u_0^\beta)(u^\gamma - u_0^\gamma) + \dots , \end{aligned} \quad (B.3)$$

where A_α^r , $\Omega_{\alpha\beta}^r$, $\Phi_{\alpha\beta\gamma}^r$, ... are coefficient sets evaluated at P . The set $\Omega_{\alpha\beta}^r$ is symmetric in α and β , the set $\Phi_{\alpha\beta\gamma}^r$ is symmetric in α , β , and γ , etc. The differentiation of (B.3) yields

$$\partial x^r / \partial u^\epsilon = A_\alpha^r \delta_\epsilon^\alpha + (1/2) \Omega_{\alpha\beta}^r [\delta_\epsilon^\alpha (u^\beta - u_0^\beta) + (u^\alpha - u_0^\alpha) \delta_\epsilon^\beta] + \dots ,$$

which, upon taking advantage of the symmetry, becomes

$$\partial x^r / \partial u^\alpha = A_\alpha^r + \Omega_{\alpha\beta}^r (u^\beta - u_0^\beta) + (1/2) \Phi_{\alpha\beta\gamma}^r (u^\beta - u_0^\beta) (u^\gamma - u_0^\gamma) + \dots \quad (B.4)$$

Similarly, we form

$$\partial^2 x^r / \partial u^\alpha \partial u^\beta = \Omega_{\alpha\beta}^r + \Phi_{\alpha\beta\gamma}^r (u^\gamma - u_0^\gamma) + \dots \quad (B.5)$$

$$\partial^3 x^r / \partial u^\alpha \partial u^\beta \partial u^\gamma = \Phi_{\alpha\beta\gamma}^r + \dots \quad (B.6)$$

etc., where the dots indicate higher-order terms.

At the point P, where $u^\alpha = u_0^\alpha$, we have from (B.3)-(B.6), respectively:

$$x^r = x_0^r; \quad (B.3')$$

$$\partial x^r / \partial u^\alpha = A_\alpha^r, \quad (B.4')$$

$$\partial^2 x^r / \partial u^\alpha \partial u^\beta = \Omega_{\alpha\beta}^r, \quad (B.5')$$

$$\partial^3 x^r / \partial u^\alpha \partial u^\beta \partial u^\gamma = \Phi_{\alpha\beta\gamma}^r. \quad (B.6')$$

etc., confirming the coefficient sets in the Taylor series expansion (B.3).

Section 3.1 has introduced A_α^r as the "design tensor", expressible via the components of u orthonormal vectors ℓ, j, \dots spanning the hyperplane tangent to the model surface at P:

$$\partial x^r / \partial u^\alpha = A_\alpha^r = \ell^\alpha \ell_\alpha^r + j^\alpha j_\alpha^r + \dots \quad (B.7)$$

which transforms like a space tensor in the contravariant indices, and like a surface tensor in the covariant indices. Any set of orthonormal vectors spanning this hyperplane is acceptable. In [Hotine, 1969] the tensor $\partial x^r / \partial u^\alpha$ was presented for $n=3$ and $u=2$. On the other hand, $\Omega_{\alpha\beta}^r, \Phi_{\alpha\beta\gamma}^r, \dots$ are not tensors. For example, from equation 6.14 in [Hotine, 1969] it follows that

$$\partial^2 x^r / \partial u^\alpha \partial u^\beta = A_{\alpha\beta}^r - \Gamma_{st}^r A_\alpha^t A_\beta^s + \Gamma_{\alpha\beta}^\gamma A_\gamma^r,$$

where $A_{\alpha\beta}^r$, the surface covariant derivative of A_α^r with respect to u^β , is a tensor, but $\partial^2 x^r / \partial u^\alpha \partial u^\beta = \Omega_{\alpha\beta}^r$ is not, due to the Christoffel symbols.

At the initial point P, a set of n orthonormal vectors $\ell, j, \dots, \nu, \dots$ spans the n -dimensional observational space. A subset of u vectors, ℓ, j, \dots , spans the above hyperplane, while the remaining $n-u$ vectors, ν, \dots , are normal to it. The observational-space metric tensor at P is expressible by

$$g_{sr} = \ell_s \ell_r + j_s j_r + \dots + \nu_s \nu_r + \dots \quad (B.8)$$

The model-surface metric tensor at P is

$$a_{\beta\alpha} = \ell_\beta \ell_\alpha + j_\beta j_\alpha + \dots = A_\beta^s g_{sr} A_\alpha^r, \quad (B.9)$$

which follows from (B.7) and (B.8). In matrix notation, $a_{\beta\alpha}$ would be represented by a positive-definite matrix of dimensions $(u \times u)$.

Next, consider another u -dimensional surface embedded in the observational space and containing the point P. The independent variables $\{u^\alpha\}$, $\alpha=1,2,\dots,u$ are again adopted to be the surface coordinates, and their values at P are again u_0^α . The spatial description of the new surface is provided by

$$x'^r = x'^r(u^\alpha) = x_0^r + A_\alpha^r (u^\alpha - u_0^\alpha). \quad (B.10)$$

Contrasted to (B.4)-(B.6), the differentiation now yields

$$\partial x'^r / \partial u^\alpha = A_\alpha^r \equiv \text{constant}, \quad (B.11)$$

$$\partial^2 x'^r / \partial u^\alpha \partial u^\beta \equiv 0, \quad \partial^3 x'^r / \partial u^\alpha \partial u^\beta \partial u^\gamma \equiv 0, \quad (B.12a,b)$$

etc. Similar to (B.3') and (B.4'), at P we have $x'^r = x_0^r$ and $\partial x'^r / \partial u^\alpha = A_\alpha^r$, while all of $\Omega_{\alpha\beta}^r$, $\Phi_{\alpha\beta\gamma}^r$, ... are zero. Since A_α^r in (B.11) is constant, and g_{sr} is constant by definition, the metric tensor for the new surface is

$$a_{\beta\alpha} = A_\beta^s g_{sr} A_\alpha^r \equiv \text{constant}, \quad (B.13)$$

implying that the surface is flat. It is thus a hyperplane, which we call the "model plane". Necessarily, the hyperplane tangent to the model plane at P is the model plane itself.

In analogy to (B.7), the constant tensor in (B.11) is written in the form

$$\partial x'^r / \partial u^\alpha = A_\alpha^r = \ell'^r \ell'_\alpha + j'^r j'_\alpha + \dots, \quad (B.14)$$

where the orthonormal vectors ℓ' , j' , ... lie entirely in the model plane. Since the tensors in (B.7) and (B.14) are equal, we anticipate that the vector sets ℓ , j , ... and ℓ' , j' , ... can be made identical, and thus that the model plane is tangent to the model surface at P. This follows from the facts below:

1) The independent variables $\{u^\alpha\}$ represent the coordinate system for both the model surface and the model plane;

2) The model-surface metric tensor $a_{\beta\alpha}$ at P is the same as its (constant) counterpart for the model plane; and

3) The model-surface design tensor A_{α}^r at P is the same as its (constant) counterpart for the model plane.

In carrying out the proof, we let the coordinate differences du^{α} represent an infinitesimal model-plane displacement from P along the unit vector ℓ' . The contravariant components of ℓ' in model-plane coordinates are $\ell'^{\alpha} = du^{\alpha}/ds$, where ds is the length of the displacement obtained from $ds^2 = du^{\beta} a_{\beta\alpha} du^{\alpha}$. Similarly, one formulates j'^{α} , etc., where

$$\ell'^{\alpha} \ell'_{\alpha} = 1, \quad \ell'^{\alpha} j'_{\alpha} = 0, \quad \dots$$

Due to item 1, the set du^{α} can also represent an infinitesimal displacement in the model surface. We denote the length of this displacement by $d\ell$, and a unit vector in the direction of the displacement by ℓ . The contravariant components of ℓ in model-surface coordinates are $\ell^{\alpha} = du^{\alpha}/d\ell$. Due to item 2, we have $d\ell = ds$ and thus $\ell^{\alpha} = \ell'^{\alpha}$. In the same way, $j^{\alpha} = j'^{\alpha}$, Due again to item 2, we have $\ell_{\alpha} = \ell'_{\alpha}$, $j_{\alpha} = j'_{\alpha}$, ... , so that

$$\ell^{\alpha} \ell_{\alpha} = 1, \quad \ell^{\alpha} j_{\alpha} = 0, \quad \dots$$

Accordingly, the thus constructed vectors ℓ , j , ... behave like model-surface orthonormal vectors at P.

With regard to the spatial configuration of ℓ , j , ... , item 3 reads

$$\ell^r \ell_{\alpha} + j^r j_{\alpha} + \dots = \ell'^r \ell'_{\alpha} + j'^r j'_{\alpha} + \dots,$$

where (B.7) and (B.14) have been utilized together with the fact that $\ell'_{\alpha} = \ell_{\alpha}$, $j'_{\alpha} = j_{\alpha}$, Upon successive contractions with ℓ^{α} , j^{α} , ... , it follows that

$$\ell^r = \ell'^r, \quad j^r = j'^r, \quad \dots \quad (B.15)$$

From the equality of their space components we deduce that the vectors ℓ and ℓ' , j and j' , etc., are identical. Since ℓ , j , ... are tangent to the model surface at P, and ℓ' , j' , ... lie entirely in the model plane, the model plane defined by (B.10) is seen to be tangent to the model surface defined by (B.3). The point of tangency, P, is described by u_0^{α} in the common surface coordinate system $\{u^{\alpha}\}$, and by $x_0^r = x^r(u_0^{\alpha})$ in the space coordinate system $\{x^r\}$.

APPENDIX C

LINKS BETWEEN OBSERVATIONAL-SPACE, MODEL-PLANE, AND MODEL-SURFACE COORDINATE DIFFERENCES

If, as in the current study, the underlying observational space is flat and the space coordinate system $\{x^r\}$ is such that

$$g_{sr} \equiv \text{constant} ,$$

in analogy to Cartesian coordinates it can be shown that

$$\Delta x^r = \Delta \rho^r , \quad (C.1)$$

where the set Δx^r contains the observational-space coordinate differences between two points, and $\Delta \rho^r$ symbolizes the difference, in contravariant components, between the position vectors of these two points. The set $\Delta \rho^r$ gives rise to the tensor equation

$$\Delta \rho^r = \Delta s \, k^r , \quad (C.2)$$

where k^r represents the contravariant components of the unit vector k in the direction of $\Delta \rho$, and Δs is the length of $\Delta \rho$. Just as in Cartesian coordinates, the position vector ρ belonging to any point in space can be freely parallel-transported from that point to an arbitrary location, and the same can be said about $\Delta \rho$ and k .

Since we shall use exclusively $\{x^r\}$ as the observational-space coordinate system, we write Δx^r for $\Delta \rho^r$, although in general coordinates (C.1) would not hold true (it is not a tensor equation). From (C.2) we deduce

$$\Delta s = \Delta x^r k_r = (1/\Delta s) \, \Delta x^r \, \Delta x_r ,$$

or

$$\Delta s^2 = \Delta x^s \, \Delta x_s = \Delta x^s \, g_{sr} \, \Delta x^r , \quad (C.3)$$

where

$$\Delta x_s = g_{sr} \, \Delta x^r . \quad (C.4)$$

Equation (C.4) is written in lieu of the tensor equation $\Delta \rho_s = g_{sr} \, \Delta \rho^r$.

Similar considerations apply for the model plane with $a_{\beta\alpha}$ constant as given by (B.13). In analogy to (C.1) and (C.2), we thus write

$$\Delta u^\alpha = \Delta \rho^\alpha, \quad (C.5)$$

$$\Delta \rho^\alpha = \Delta s k^\alpha. \quad (C.6)$$

Here the symbols $\Delta \rho$, Δs , and k are used independently of the same symbols appearing in (C.1,2). In analogy to (C.4), we also have

$$\Delta u_\beta = a_{\beta\alpha} \Delta u^\alpha, \quad (C.7)$$

representing the tensor equation $\Delta \rho_\beta = a_{\beta\alpha} \Delta \rho^\alpha$.

In linking observational-space and model-surface coordinate differences we shall proceed via the model plane, where we can take advantage of the property (C.5). Let $\Delta x'$ represent a model-plane vector emanating from P, which, in Fig. 1, is the vector symbolized by the arrow PP_2 . Upon considering ℓ, j, \dots to be orthonormal vectors spanning the model plane, it follows that

$$\Delta x'^r = a \ell^r + b j^r + \dots, \quad (C.8)$$

where a, b, \dots are scalars. If the same vector should be described by model-plane rather than observational-space components, we denote it Δu and write

$$\Delta u^\alpha = a \ell^\alpha + b j^\alpha + \dots \quad (C.9)$$

Due to (B.14), presented here as

$$A_\alpha^r = \ell^r \ell_\alpha + j^r j_\alpha + \dots, \quad (C.10)$$

one obtains

$$\Delta x'^r = A_\alpha^r \Delta u^\alpha, \quad (C.11)$$

where $\Delta x'^r$ symbolizes the contravariant components of the above vector as well as the corresponding coordinate differences in the observational-space coordinates, and Δu^α symbolizes such components and such coordinate differences in the model-plane coordinates.

Since both (C.8) and (C.9) hold true also when the indices are lowered, we readily deduce that

$$\Delta u_\beta = A_\beta^s g_{sr} \Delta x'^r. \quad (C.12)$$

This relation can also be obtained from (C.11), upon contracting the latter with $A_{\beta}^s g_{sr}$ and recalling (C.7) together with (B.13). As a special benefit, if $\Delta x'$ were replaced by a general space vector Δx , the left-hand side of (C.12) would describe a projection of that vector onto the model plane.

To confirm this property, suppose that the contravariant components of a space vector are $\Delta x'^r + \Delta x''^r$, where

$$\Delta x''^r = q \nu^r + \dots$$

In terms of Fig. 1, we can imagine $\Delta x''$ as a vector perpendicular to the model plane at P_2 . It then follows that

$$\Delta u_{\beta} = A_{\beta}^s g_{sr} (\Delta x'^r + \Delta x''^r) = A_{\beta}^s g_{sr} \Delta x'^r. \quad (C.13)$$

In other words, if the components $\Delta x'^r + \Delta x''^r$ are known, (C.13) yields the model-plane covariant components of the projected vector, Δu_{β} . The observational-space contravariant components of the projected vector, $\Delta x'^r$, can be obtained via (C.11) upon first utilizing

$$\Delta u^{\alpha} = a^{\alpha\beta} \Delta u_{\beta}, \quad (C.14)$$

where $a^{\alpha\beta}$ is the model-plane associated metric tensor (such that $a^{\alpha\beta} a_{\beta\gamma} = \delta^{\alpha}_{\gamma}$). One feature of equations (C.13) and (C.14) is that they demonstrate how the observational-space components of a vector can be converted into model-plane coordinate differences.

We shall next present a spatial link between the model surface and the model plane by relating corresponding points in the two surfaces. In Fig. 1, a pair of such points is depicted by \bar{P} and P_2 . Their correspondence is understood in the sense that the same set of values u^{α} represents the model-surface coordinates of \bar{P} and the model-plane coordinates of P_2 . In the observational space, the points \bar{P} and P_2 are identified by their position-vector components x^r and x'^r , respectively. By subtracting from either of the latter the position-vector components of P , x_0^r , we have

$$\Delta x^r = x^r - x_0^r, \quad \Delta x'^r = x'^r - x_0^r.$$

These sets of contravariant components, which represent also sets of coordinate differences, identify the vectors Δx and $\Delta x'$ symbolized in Fig. 1 by the arrows

$P\bar{P}$ and PP_2 , respectively. Upon denoting

$$\Delta u^\alpha = u^\alpha - u_0^\alpha,$$

the comparison of (B.3) and (B.10) reveals that

$$\Delta x^r = \Delta x'^r + v^r, \quad (C.15)$$

where the contravariant components v^r of the vector v are given as

$$v^r = (1/2) \Omega_{\alpha\beta}^r \Delta u^\alpha \Delta u^\beta + (1/6) \Phi_{\alpha\beta\gamma}^r \Delta u^\alpha \Delta u^\beta \Delta u^\gamma + \dots \quad (C.16)$$

The vector v provides the desired link between the two surfaces.

In considering that the symbols x are interchangeable here with ρ , the relation (C.15) is a tensor equation. On the other hand, since $\Omega_{\alpha\beta}^r$, $\Phi_{\alpha\beta\gamma}^r$, ... are not tensors, (C.16) is not a tensor equation, but this does not detract from its usefulness. We observe that tensor considerations have been important mainly in deriving (C.13), where the symbols Δu and $\Delta x'$ represent the same geometrical object. The tensor contraction with $a^{\alpha\beta}$ in (C.14) then yields Δu^α , which identifies both the set of contravariant components of the vector PP_2 and the set of model-plane coordinate differences between P_2 and P . Due to the definition of $\{u^\alpha\}$, the set Δu^α represents also model-surface coordinate differences between \bar{P} and P .

APPENDIX D

MATRIX NOTATION

If needed, the results of this study can be transcribed into matrix notation. This is greatly facilitated by the close correspondence between tensor contractions and matrix multiplications. In the cases of $\Omega_{\alpha\beta}^r$ and $\Phi_{\alpha\beta\gamma}^r$, respectively, we introduce three- and four-dimensional arrays into the matrix context. In transcribing the corresponding quantities, the indices will be dropped. The letter symbols will be retained, but * will be attached to quantities corresponding to purely covariant tensors. Thus, the tensors g^{rs} , A_{α}^r , $a^{\alpha\beta}$, δx^r , Δx^r , Δu^{α} , etc., will become g , A , a , δx , Δx , Δu , etc., while the tensors g_{sr} , $a_{\beta\alpha}$, Δu_{β} , etc., will become g^* , a^* , Δu^* , etc. The first index identifies the rows and the second index, if present, identifies the columns. If the indices are mixed, as in A_{α}^r , the upper (contravariant) index is regarded as the first and the lower (covariant) index is regarded as the second.

The sets $\Omega_{\alpha\beta}^r$ and $\Phi_{\alpha\beta\gamma}^r$ will be transcribed respectively as a three-dimensional array Ω of dimensions $(n \times u \times u)$ and a four-dimensional array Φ of dimensions $(n \times u \times u \times u)$. In analogy to the matrix A of dimensions $(n \times u)$ evaluated at P according to

$$A \equiv [\partial x / \partial u^1, \partial x / \partial u^2, \dots] ,$$

where x represents the observables and u^1, u^2, \dots are the individual parameters, the arrays Ω and Φ are formed at P as

$$\Omega \equiv [\partial A / \partial u^1, \partial A / \partial u^2, \dots] , \quad \Phi \equiv [\partial \Omega / \partial u^1, \partial \Omega / \partial u^2, \dots] .$$

Upon considering the order of contractions, it becomes clear that the expression $A_{\beta}^s g_{sr}$ corresponds to $A^T g^*$ (and not, for example, to $A g^*$ which is not even defined). The contractions of the kind $\Omega_{\alpha\beta}^r (\Delta u^{\beta})$ and $(\delta x_s) \Phi_{\beta\alpha\gamma}^s$ will be transcribed as $\Omega(\Delta u)$ and $(\delta x^*)^T \Phi$, respectively.

In dealing with quantities related to the observables, such as x_Q^r , x^r , x_0^r , δx^r , Δx^r , $\delta \bar{x}^r$, etc., which in the geometrical context of the observational space represented sets of coordinates or sets of coordinate differences equivalent to tensors, we again drop the indices and use the symbols x_Q , x , x_0 , δx , Δx , $\delta \bar{x}$, etc. The column vectors (of n elements) x_Q , x , and x_0 denote respectively the actual observations, the adjusted observations, and the

observables consistent with the initial set of parameters. The column vectors δx , Δx , and $\delta \bar{x}$, the latter containing the negative residuals, follow respectively from the relation below (8), from (4), and from (14):

$$\delta x \equiv x_Q - x_0 ;$$

$$\Delta x = x - x_0 , \quad \delta \bar{x} \equiv x_Q - x = \delta x - \Delta x .$$

Quantities related to the parameters, such as u^α , u_0^α , Δu^α , etc., which in the geometrical context of the model plane are sets of coordinates or sets of coordinate differences equivalent to tensors, are symbolized by u , u_0 , Δu , etc. The column vectors (of u elements) u and u_0 denote respectively the adjusted parameters and the initial values of parameters, while the column vector Δu follows from (1b) as

$$\Delta u = u - u_0 .$$

Introducing also quantities in parentheses, in analogy to Δx and $\delta \bar{x}$ we have

$$(\Delta x) = (x) - x_0 , \quad (\delta x) \equiv x_Q - (x) = \delta x - (\Delta x) .$$

Moreover, in analogy to Δu we write

$$(\Delta u) = (u) - u_0 .$$

The intermediate values such as (u) and (x) follow from these equations as $(u) = u_0 + (\Delta u)$ and $(x) = x_0 + (\Delta x)$. Similar transcriptions apply also for the relationships involving $\Delta(\Delta u^\alpha)$ and $\Delta(\Delta x^T)$, such as presented in (15) and (19a), respectively, but it is unnecessary to write them explicitly since the pattern explained above is general and straightforward.

We next transcribe the key formulas associated with the initial iteration of Section 3.1. Equations (6), (8), and (9) correspond to

$$a^* = A^T g^* A , \tag{D.1}$$

$$(\Delta u^*) = A^T g^* \delta x , \quad (\Delta u) = a (\Delta u^*) , \tag{D.2a,b}$$

where a is the inverse of a^* . Further, the formulas (10a,b) become

$$g' = A a A^T , \quad g^{*'} = g^* g' g^* , \tag{D.3a,b}$$

while the relations corresponding to (11a,b) are

$$g'' = g - g' , \quad g^{*''} = g^* - g^{*'} . \quad (D.4a,b)$$

The adjustment role of the matrices g' , $g^{*'}$, g'' , and $g^{*''}$ of dimensions $(n \times n)$ has already been explained. Similar to the relationship between the matrices a and a^* of dimensions $(u \times u)$, the matrices g and g^* of dimensions $(n \times n)$ are inverses of each other.

Finally, we present the system (23a-c) in matrix notation:

$$[(a^*) - (\delta x^*)^T \Omega - (\delta x^*)^T \Phi (\Delta u)] \Delta(\Delta u) = (A)^T (\delta x^*) , \quad (D.5a)$$

where

$$(a^*) = (A)^T g^* (A) , \quad (\delta x^*) = g^* (\delta x) . \quad (D.5b,c)$$

We call (D.5a) the modified normal equations, and the matrix on the left-hand side, denoted $\{a^*\}$, the matrix of modified normal equations. For a repeated application of (D.5a-c), (Δu) is updated through the algebraic addition of $\Delta(\Delta u)$. All the other quantities in parentheses are updated as well, while the arrays Ω and Φ are treated as constant. Depending on the assumption about the size of the residuals, (D.5a) can be simplified by leaving out the term with Φ . If both terms with Ω or Φ are left out, the resulting equation characterizes the standard adjustment approach. The final updated values are denoted by overbars. The variance-covariance and the weight matrices of adjusted quantities are represented by (D.1), (D.3a,b), and (D.4a,b), with all the symbols except g and g^* overbarred, and with \bar{a} computed as the inverse of \bar{a}^* .

REFERENCES

- Besette, R. P., 1987. "Nonlinear Parametric Least Squares Adjustment -- Example". Paper presented at the AGU Fall Meeting, San Francisco, Cal.; EOS 68 (44), p. 1237.
- Blaha, G., 1984. *Analysis Linking the Tensor Structure to the Least-Squares Method*. AFGL Technical Report 84-0084, Air Force Geophysics Laboratory, Bedford, Mass.; ADA142159.
- Blaha, G., 1987. *Nonlinear Parametric Least-Squares Adjustment*. AFGL Technical Report 87-0204, Air Force Geophysics Laboratory, Bedford, Mass.; ADA184039.
- Hotine, M., 1969. *Mathematical Geodesy*. ESSA Monograph 2, U.S. Department of Commerce, Washington, D.C.
- Teunissen, P. J. G., 1985. *The Geometry of Geodetic Inverse Linear Mapping and Non-Linear Adjustment*. Vol. 8, No. 1 in Publications on Geodesy of the Netherlands Geodetic Commission, Delft.